

```

/product="ATP synthase"
/protein_id="AAB87466.1"
/db_xref="GI:2641643"
/translation="MATEKIVQIIICAVIDVEFPDQAVPRKVDALVESGLTEVQQQL
GGGLVRCIALGTSIDGLKRLKLVENGNPQVGVGTIKLRIIMNVLGEPIDEKPIGEV"

BASE COUNT      97 a      54 c      79 g      77 t

ORIGIN

Query Match      1.8%; Score 35; DB 1; Length 307;
Best Local Similarity 100.0%; Pred. No. 1.4e-07;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1308 caagatgcagtaacaaagatatatgctccttaaa 1342
|||||
Db 58 CAAGATGCAGTACCAAAAGTATATGATGCTTMAA 92

RESULT 6
AF037146      793 bp      DNA      BCT      19-MAY-1998
LOCUS      Salmoneilla enterica strain S83769 ATP synthase beta subunit (atpD)
DEFINITION      gene, partial cds.
ACCESSION      AF037146
VERSION      AF037146.1 GI:3138947
KEYWORDS
SOURCE      Salmoneilla enterica.
ORGANISM      Salmoneilla enterica.
FEATURES
    source
        Salmoneilla enterica.
        Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
        Salmoneilla.
    REFERENCE
        1 (bases 1 to 793)
        Christensen, H. and Olsen, J.E.
        Phylogenetic relationships of Salmoneilla based on DNA sequence
        comparison of atpD encoding the beta subunit of ATP synthase
        FEMS Microbiol. Lett. 161 (1), 89-96 (1998)
    JOURNAL
        MEDLINE
        98222522
    REFERENCE
        2 (bases 1 to 793)
        Christensen, H. and Olsen, J.E.
        Direct Submission
        Submitted (08-DEC-1997) Veterinary Microbiology, the Royal
        Veterinary and Agricultural University, Bulowstvej 13, Copenhagen
        1870 Fred. C., Denmark
    FEATURES
        Location/Qualifiers
            1..793
            /organism="Salmoneilla enterica"
            /strain="S83769"
            /serotype="T11a 51:z4,z23:-"
            /db_xref="taxon:28901"
            <1..>793
            /gene="atpD"
            <1..523
            /gene="atpD"
            /codon_start=2
            /transl_table=11
            /product="ATP synthase beta subunit"
            /protein_id="AAC16682.1"
            /db_xref="GI:3138948"
            /translation="NGNEKLVLEVOOOLGGIVRTIAMSSDGLRGLDVKDLHPHIE
            VPKATIGRIMNVLGEPIVDMKGEIPEERMAIHRAAPSEELNSOELLEFGIVYID
            LMCPRKAGKVGELFGAGVGVKTVNMELIRNIAIHSIGSVAGVGERIREGNDPIHE
            MTDNSVIDKVSILV"

BASE COUNT      181 a      197 c      239 g      176 t

ORIGIN

Query Match      1.6%; Score 32; DB 1; Length 793;
Best Local Similarity 100.0%; Pred. No. 6.8e-06;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1920 atggcgaaaaaattccgtatgaagtcgtga 1951
|||||
Db 581 ATGGCGAAAAAATTCGATGATGAAGTCGTGA 612

```

```

RESULT 7
AF037147      793 bp      DNA      BCT      19-MAY-1998
LOCUS      Salmoneilla enterica strain u24 ATP synthase beta subunit (atpD)
DEFINITION      gene, partial cds.
ACCESSION      AF037147
VERSION      AF037147.1 GI:3138949
KEYWORDS
SOURCE      Salmoneilla enterica.
ORGANISM      Salmoneilla enterica.
FEATURES
    source
        Salmoneilla enterica.
        Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
        Salmoneilla.
    REFERENCE
        1 (bases 1 to 793)
        Christensen, H. and Olsen, J.E.
        Phylogenetic relationships of Salmoneilla based on DNA sequence
        comparison of atpD encoding the beta subunit of ATP synthase
        FEMS Microbiol. Lett. 161 (1), 89-96 (1998)
    JOURNAL
        MEDLINE
        98222522
    REFERENCE
        2 (bases 1 to 793)
        Christensen, H. and Olsen, J.E.
        Direct Submission
        Submitted (08-DEC-1997) Veterinary Microbiology, the Royal
        Veterinary and Agricultural University, Bulowstvej 13, Copenhagen
        1870 Fred. C., Denmark
    FEATURES
        Location/Qualifiers
            1..793
            /organism="Salmoneilla enterica"
            /strain="u24"
            /serotype="T11a [Arizonae]"
            /db_xref="taxon:28901"
            <1..>793
            /gene="atpD"
            <1..>793
            /gene="atpD"
            /codon_start=2
            /transl_table=11
            /product="ATP synthase beta subunit"
            /protein_id="AAC16683.1"
            /db_xref="GI:3138950"
            /translation="NGNEKLVLEVOOOLGGIVRTIAMSSDGLRGLDVKDLHPHIE
            VPKATIGRIMNVLGEPIVDMKGEIPEERMAIHRAAPSEELNSOELLEFGIVYID
            LMCPRKAGKVGELFGAGVGVKTVNMELIRNIAIHSIGSVAGVGERIREGNDPIHE
            MTDNSVIDKVSILVQGMNEPGRNRLVALTGLMAKEFDEBDDVLLFYDNIYRTTLA
            GTEVSLALGRMSAVGVOPTLAEMGVLOERITSTRTGTSVOAV"

BASE COUNT      182 a      195 c      238 g      178 t

ORIGIN

Query Match      1.6%; Score 32; DB 1; Length 793;
Best Local Similarity 100.0%; Pred. No. 6.8e-06;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1920 atggcgaaaaaattccgtatgaagtcgtga 1951
|||||
Db 581 ATGGCGAAAAAATTCGATGATGAAGTCGTGA 612

RESULT 8
AF037148      793 bp      DNA      BCT      19-MAY-1998
LOCUS      Salmoneilla enterica strain JEO307 ATP synthase beta subunit (atpD)
DEFINITION      gene, partial cds.
ACCESSION      AF037148
VERSION      AF037148.1 GI:3138951
KEYWORDS
SOURCE      Salmoneilla enterica.
ORGANISM      Salmoneilla enterica.
FEATURES
    source
        Salmoneilla enterica.
        Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
        Salmoneilla.
    REFERENCE
        1 (bases 1 to 793)
        Christensen, H. and Olsen, J.E.

```

TITLE	Phylogenetic relationships of salmonella based on DNA sequence comparison of atpD encoding the beta subunit of ATP synthase
JOURNAL	FEMS Microbiol. Lett. 161 (1), 89-96 (1998)
MEDLINE	98222522
REFERENCE	2 (bases 1 to 793)
AUTHORS	Christensen, H. and Olsen, J.E.
TITLE	Direct Submission
JOURNAL	Submitted (08-DEC-1997) Veterinary Microbiology, the Royal Veterinary and Agricultural University, Bulowsvej 13, Copenhagen 1870 Fred. C., Denmark
FEATURES	location/Qualifiers
source	1..793
gene	/organism="Salmonella enterica" /strain="JE0307" /serotype="11:b:6:1:2" /db_xref="taxon:28901" <1..>793 /gene="atpD" <1..>793 /gene="atpD" /codon_start=2 /product="ATP synthase beta subunit" /protein_id="A0C16684.1" /db_xref="GI:3138952" /translation="NGNEKLVLEVOOALGGIVRTTAMGSSDGLRGLDVKDEHPLEVPKATIGRIIMNVLEPEPMVKGEICIGEEFWALHRRAPSYEELNSOELLETGKIVDLMCPKAGVGLFEGAGVGVKVTNMMLINIMAIHESGVSVFAGVGSGTEGNDFYHEMTSPNIDKVSIVYGGNNEPPGNRLIRVALGIMPAEKFRDEGRDVLVLFVNIYKYTLACTEVSALGKMPASVGVQPTLAEPMGLQBRISTKTSGLTSYQAV"
CDS	182 a 194 c 238 g 179 t
BASE COUNT	
ORIGIN	

```

Query Match          1.6%; Score 32; DB 1; Length 793;
Best Local Similarity 100.0%; Pred. No. 6 8e-06;
Matches      32; Conservative    0; Mismatches     0; Indels       0; Gaps
Oy      1920 atgcggaataatccgtgatgaagtcgtga 1951
|||||
|||ATGGCGGAAAAATTCCGTGATGAAGCTGCTGA 612
Db      581

RESULT      9
AF037149      793 bp      DNA      BCT      19-MAY-1998
LOCUS      Salmomella enterica strain DS210/89 ATP synthase beta subunit
DEFINITION      (acpd) gene, partial cds.
ACCESSION      AF037149
VERSION      AF037149
KEYWORDS      AF037149.1 GI:3138953
SOURCE      .
ORGANISM      Salmomella enterica.
Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
Salmomella.
1 (bases 1 to 793)
Christensen,H. and Olsen,J.E.
Phylogenetic relationships of Salmomella based on DNA sequence
comparison of acpd encoding the beta subunit of ATP synthase
FEMS Microbiol. Lett. 161 (1), 89-96 (1998)
98222522
2 (bases 1 to 793)
Christensen,H. and Olsen,J.E.
Direct Submission
Submitted (08-DEC-1997) Veterinary Microbiology, the Royal
Veterinary and Agricultural University, Bulowsevej 13, Copenhagen
1870 Fred. C., Denmark
Location/Qualifiers
1..793
/organism="Salmomella enterica"
/strain="DS210/89"
/serotype="Tiib 48:r:z"
/db_xref="taxon:28901"
FEATURES
source
```

```

gene
CDS
<1..>793
/feature="atpD"
<1..>793
/feature="atpD"
/codon_start=2
/transl_table=11
/product="ATP synthase beta subunit"
/protein_id="AAC16685.1"
/db_xref="GI:3138954"
/translation="NGNEKLVLEVOOQLGGGIVRTIAMSSDGLRLGLDYKDEHPLE
VPPKQKATIGIRIMNVLPGEVDMKGEIGSEERMAIRAPSYEELSNSOELLETKGKVID
LMKPFPAKGVGLFGGAGCYCKVTMMELINIAIEHSGSVFAVGEGTREGNGFYHE
MTQSNVIDKSLVLYGGQNEPPGNRLRLRVALIGLITAAEKFRDEGRVLLFVYNIYYTLA
GTEVSALIGRMPsAVGYOPTLAEEMGLQERITSTKGSITSVOAV"

```

Query Match	1.6%	Score 32;	DB 1;	Length 793;
Best Local Similarity	100.0%	Pred. No. 6	8e-06;	
Matches	32;	Conservative	0;	Mismatches 0; Indels 0; Gaps 0;
Q1	1920	atgcgcggaataatccgtgatgaagatcgtga	1951	
Db	581	ATGCGCGAATAATTCGCGATGAGGTCGTGA	612	
RESULT	10			
AF037150				
LOCUS				
DEFINITION	AF037150	793 bp	DNA	BCT 19-MAY-1998
ACCESSION	AF037150			
VERSION	AF037150			
KEYWORDS	AF037150.1	GI:3138955		
SOURCE				
ORGANISM				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
MEDLINE				
LOCUS				
AUTHORS				
TITLE				
JOURNAL				
MEDLINE				
LOCUS				
AUTHORS				
TITLE				
JOURNAL				
MEDLINE				
LOCUS				
AUTHORS				
TITLE				
JOURNAL				
MEDLINE				
LOCUS				
AUTHORS				
TITLE				
JOURNAL				
MEDLINE				
LOCUS				
AUTHORS				
TITLE				
JOURNAL				
MEDLINE				
LOCUS				
AUTHORS				
TITLE				
JOURNAL				
MEDLINE				
LOCUS				
AUTHORS				
TITLE				
JOURNAL				
MEDLINE				
LOCUS				
AUTHORS				
TITLE				
JOURNAL				
MEDLINE				
LOCUS				
AUTHORS				
TITLE				
JOURNAL				
MEDLINE				
LOCUS				
AUTHORS				
TITLE				
JOURNAL				
MEDLINE				
LOCUS				
AUTHORS				
TITLE				
JOURNAL				
MEDLINE				
LOCUS				
AUTHORS				
TITLE				
JOURNAL				
MEDLINE				
LOCUS				
AUTHORS				
TITLE				
JOURNAL				
MEDLINE				
LOCUS				
AUTHORS				
TITLE				
JOURNAL				
MEDLINE				
LOCUS				
AUTHORS				
TITLE				
JOURNAL				
MEDLINE				
LOCUS				
AUTHORS				
TITLE				
JOURNAL				
MEDLINE				
LOCUS				
AUTHORS				
TITLE				
JOURNAL				
MEDLINE				
LOCUS				
AUTHORS				
TITLE				
JOURNAL				
MEDLINE				
LOCUS				
AUTHORS				
TITLE				
JOURNAL				
MEDLINE				
LOCUS				
AUTHORS				
TITLE				
JOURNAL				
MEDLINE				
LOCUS				
AUTHORS				
TITLE				
JOURNAL				
MEDLINE				
LOCUS				
AUTHORS				
TITLE				
JOURNAL				
MEDLINE				
LOCUS				
AUTHORS				
TITLE				
JOURNAL				
MEDLINE				
LOCUS				
AUTHORS				
TITLE				
JOURNAL				
MEDLINE				
LOCUS				
AUTHORS				
TITLE				
JOURNAL				
MEDLINE				
LOCUS				
AUTHORS				

Query Match 1.6%; Score 32; DB 1; Length 793;
Best Local Similarity 100.0%; Pred. No. 6.8e-06;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1920 atggcgaaaattccgtatgaagtcgtga 1951
|||||
DB 581 ATGGCGAAAAATTCGTCATGAAGCTCGTGA 612

RESULT 11
LOCUS BACATPA 7307 bp DNA BCT 29-NOV-1995
DEFINITION B.megaterium ATP synthase i,a,c,b,delta,alpha,gamma,beta and
epsilon subunit genes, complete cds, and ORF.
ACCESSION M20255 J04455 M18352 M23924
VERSION M20255.1 GI:142553
KEYWORDS ATP synthase; ATP synthase I subunit; ATP synthase alpha subunit;
epsilon subunit; ATP synthase gamma-subunit; ATPase; ORF.
SOURCE B.megaterium (QM B151) DNA, clones pWSB100, pCAH1.3, and pWPC208.
ORGANISM Bacillus megaterium
Bacteria; Firmicutes; Bacillus/Clostridium group;
Bacillus/Staphylococcus group; Bacillus.
REFERENCE 1 (bases 5401 to 7307)
Hawthorne,C.A. and Brusilow,W.S.
Sequence of the genes for the beta and epsilon subunits of the ATP
synthase of *Bacillus megaterium* QM B151
Biochem. Biophys. Res. Commun. 151: (2), 926-931. (1988)
JOURNAL 198162920
AUTHORS 2 (bases 1 to 5400)
TITLE Brusilow,W.S., Scarpetta,M.A., Hawthorne,C.A. and Clark,W.P.
Organization and sequence of the genes coding for the
proton-translocating ATPase of *Bacillus megaterium*
J. Biol. Chem. 264 (3), 1528-1533 (1989)
JOURNAL 89109162
COMMENT Draft entry and computer-readable sequence for [2] kindly submitted
by W.S.A. Brusilow, 09-NOV-1988.
FEATURES
source
1. 7307
Location/Qualifiers
/organism="Bacillus megaterium"
/db_xref="taxon:1404"
1..264
/note="ORF (AA at 1); putative"
/codon_start=1
/transl_table=1
/product="unknown protein"
/protein_id="AA82518.1"
/db_xref="GI:1088271"
/translation="EFGKKILQKYEELMKDRNRHPLQAMALVSIILSOLAGSVLIGL
FSGRADQMLDEPLFLILGLGLAAGFMALRTVHHFFSD"
271..642
/codon_start=1
/transl_table=1
/product="ATP synthase i subunit"
/protein_id="AA82519.1"
/db_xref="GI:142554"
/translation="MODLOHVPRLRSYLLALAYVLGCGFTSYKAVFAGLILGTAL
SLYLMNLRKEFGOALDEGKKRPSIGTVRFATALAVYITISYPTFHISIVV
GLMTYVVIIDLVQNNRRR"
661..1371
/codon_start=1
/transl_table=1
/product="ATP synthase a subunit"
/protein_id="AA82520.1"
/db_xref="GI:142555"
/translation="MGHESPTIEFLGLPFSOSNLAMVYVAVIPLAVLCRTITLAK
PSGACNTEMLVDLVKGLNSNMKRGRLTLGVTLMTLYVSNMGLPFAIVIDH
NLWKSPTADPAITLTLAVVNVLSHYGIGKMGFSAYTKDYFKPMALFLPKITIEEF
ANTLTKRLYGNITAGETILSLAGLATTTGTLTGAIAPMLLMQGSIEFGAIOAF
IFTMLTMYLSHKVSSDH"
1462..1674

CDS
/codon_start=1
/transl_table=1
/product="ATP synthase c subunit"
/protein_id="AA82521.1"
/db_xref="GI:142556"
/translation="MGLIASAIGLALGAGIGNGLIYSKTIETGARPEARGLTIS
MFEVGAIVEALPIAVIAFMVQCK"
1777..2295
/codon_start=1
/transl_table=1
/product="ATP synthase b subunit"
/protein_id="AA82522.1"
/db_xref="GI:142557"
/translation="MAVSNMFYGAAGINGDILRLQVFLILALLQKFAFGPWGI
MKRREHIGEIDEKQNEAKKIVVEERELIKSROFQVQMMENAKSADKKEI
VAAARESERLKAANKQETEDQKDAVALRQVAVSLVLSKVIETKELSDQDKL
IHETIOEVGDVR"
2292..2837
/codon_start=1
/transl_table=1
/product="ATP synthase delta subunit"
/protein_id="AA82523.1"
/db_xref="GI:142558"
/translation="MSQPAVKARVALALPOLATEKOMIDEMODLOIIVEFPAKPEL
MDVLRPKITIERKOFSEARFARLSPVQHVLLLEHRIQIYSEMYKEVRLANE
VAGVADATYVSKPLSADKRAISQSFASKVGHKHLNLSINYDKYIGVKRLIGMRI
YGSISSKLETTHRGHLAHS"
2853..4361
/codon_start=1
/transl_table=1
/product="ATP synthase alpha subunit"
/protein_id="AA82524.1"
/db_xref="GI:142559"
/translation="MSIKAEISALIKQOIEYOSIEIKYSDVGVYIOWDGIARAHL
DNVMAGEIVFSGVMAQNLKENNVGIIIGPYTEIEGDEVRTRGRIMEVPGEQ
LIGRVYNSGQVVDGIGPEYETKTRPIEGAGVMDRKSVHEPLDTGKATDALVPIG
RGQRELIIGDROTKSVADIPILNOKQDMQCIYVAGKESYRNVETLRKGLAL
DTIIVYASASQAPLLPLAPYAGVYMGEEYNGKHVYITDITLKQASATRELSL
LRRPGREAYPGDVFYILHSRLERAKISDAGGSLVAPLETQAGVSAVITPNV
ISITDGOITLQDLEFFSGVRAINAGLSVRGSGQIKAMKVGATIDIASREL
ESFAQFGSDLDQATOAKLNRGARVETILKQGLHKLRYKQAVAVYALTKGLDQVY
SDITRPEDEYLTWLESNREVEISIRTTGCLPEAGLFESALEPFTKTFIASE"
4467..5324
/codon_start=1
/transl_table=1
/product="ATP synthase gamma subunit"
/protein_id="AA82525.1"
/db_xref="GI:142560"
/translation="NASLSDIOTRITTSFKTSQITRKAMENYSAKLINRAEQNAKSFV
IWEKIOEVVSSVALSGASHPMLTARSKKGIVYITSDRLAAGVNSNLIKRSQA
IEERHQSDEYVIAIGRVGRDFPVKRGIPVLETTGGLADQVAFDIOGASQVOMF
ADGTFDELYLYNHFINTISQVETKRLPLDLIDPLSGVLEVEEPESEELTEVLLP
QYAESLIYGLIDGKASEHARMTAMK SATDNKDLINMLTISNARQAATQETITE
IVGGAALAE"
5361..6779
/codon_start=1
/transl_table=1
/product="ATP synthase beta subunit"
/protein_id="AA82526.1"
/db_xref="GI:142561"
/translation="MTKGVYQIMGPYVDVFKDNGHLPATYNAKLISHKPSASVAVI
ELTLEVAHLGDNVRYVAMSSTDLVRLVEEDGALISVPGVDTLGRVNVJGK
IDLDAVIDAGARRDPIHRQAPKFNLSQAEILLEGIRKVDLLAPYIGKGLIGLGA
GVKATVLIQELINNIQAEHGGISVAGVETRENDYIHEMTDSGVIRKTMVFGGA
NEPGARQORVOLTGLMAEYFRDEQGVLEFIDINIFRTOGASEVALGTPMPSAG
YQPLATEMGOLERTSTSVGSVSIOAIYVPADYDPPATTFALHDAITNARK
LSEKTIYPAVDPLASTSRSLSPETIYGEHEATARAQVOTIOYKELDITATIGDEL
SDEKLVVPRARVQFFLSQTSVADQFTGQKSYVAVETVAGFEILLEGKIDHLPED
ARLVRLGLEEVLENKRGVEV"
6806..7210
/codon_start=1
/transl_table=1
/product="ATP synthase epsilon subunit"

BASE COUNT 2274 a 1331 c 1639 g 2063 t
ORIGIN 1 bp upstream of EcoRI site.

Query Match 1.38; Score 26; DB 2; Length 7307;
Best Local Similarity 100.0%; Pred. No. 0.015; Indels 0; Gaps 0;
Matches 26; Conservative 0; Mismatches 0;

QY 1779 gtagctagcgtacgcgtgaagttaa 1804
|||||
DB 5922 GTAGGTAGCGGTACGCTGAAGCTAA 5947

RESULT 12
LOCUS AF008210/c
DEFINITION Buchnera aphidicola genomic fragment containing (chaperone Hsp60)
grodL, DNA biosynthesis initiating protein (dnaa), ATP operon
(agpGDAHEB), and putative chromosome replication protein (gidA)
genes, complete cds; and termination factor Rho (rho) gene, partial
cds.
ACCESSION AF008210
VERSION AF008210.1 GI:2827005
KEYWORDS Buchnera aphidicola.
SOURCE Buchnera aphidicola
ORGANISM Bacteria; Proteobacteria; gamma subdivision; Buchnera.
REFERENCE 1 (bases 1 to 34699)
AUTHORS Clark, M.A., Baumann, L. and Baumann, P.
TITLE Sequence analysis of a 34.7-kb DNA segment from the genome of
Buchnera aphidicola (endosymbiont of aphids) containing grodL,
dnaa, the atp operon, gidA, and rho
JOURNAL Curr. Microbiol. 36 (3), 158-163 (1998)
MEDLINE 98184963
JOURNAL 2 (bases 1 to 34699)
REFERENCE Clark, M.A., Baumann, L. and Baumann, P.
AUTHORS Direct Submission
TITLE Submitted (13-JUN-1997) Microbiology Section, Univ. of Calif. at
JOURNAL Davis, CA 95616-8665, USA
FEATURES
Source Location/Qualifiers
1. 34699
/organism="Buchnera aphidicola"
/specific_host="Schizaphis graminum"
/db_xref="taxon:9"
<1.437
/codon_start=3
/transl_table=1
/product="39-kDa hypothetical protein"
/protein_id="AAC38098.1"
/db_xref="GI:2827006"
/translation="MHTRLPIVPIPNRTSDLCQIFSNISVLIIVTHINHPQELNEQL
SDSLKTLKSNVILNLSVLKININDNAIIIAELSSRCENNIIPYLIHLIDKVGYS
HEIIVSNKAKSIIDLMKISGLVPLRVFNGSKDKMLIIT"
complement(558..2213)
/gene="grodL"
complement(558..2213)
/gene="grodL"
/codon_start=1
/transl_table=1
/product="chaperone Hsp60"
/protein_id="AAC38099.1"
/db_xref="GI:2827007"
/translation="MSKMGADVKFGFNGEARTIKMLGVNLADAVKVTLLGPGRNVLID
KSPGAPSTIKGVSAAREIELEDFENMGAGQVKEVASKANDAGDGTITALLAQSI
VNEGLKATAAGAMPDLKRGIDKAVISAVEELKNLSVPCSDSKAITGVGTISANDEK
VGLALAEAMERYGNDGVITVEEGTGLQMLELYVKGMPQDRGYSYFINKREGIVEL
ENPYILMDKRIISNREMLPILESVAKSGKPLIITSEDLGEBALATLVNPNRGIIVY
AAVAPGFGDRRKAMLDISITLTGGSVISSEELAMDLEKSTLEDLGQAKRVVINKDTT

ITGSGSKOAIOSRIGOIROEIOEATSDPKKEKLEIRLAKISGVAVLKVGATEVEN
KKKARVBDALATRAAVEGVAGGVALVAVACKISLNRGHNDDQVGLVIALRAM
EAPLKOIYSNCSSEBSVYTNKKDKDKGNGVAAADDEGMDIDPGLDPTKYTRBALQ
YAASVAGMLTTECEWYTDLPREDKSDVAASPAAGMGNGMGAM"
complement(2262..2552)
/gene="groES"
complement(2262..2552)
/gene="groES"
/transl_table=1
/codon_start=1
/product="chaperone Hsp10"
/protein_id="AAC38100.1"
/db_xref="GI:2827008"
/translation="MKIRPLHRYLVKRNBAELKSGAVLVNGSAAKSTRGITAVG
NGRVLDNGOIKPLDVAVGQVYIFNEGYGAKTEKIDNEELLINSDILAVE"
complement(2795..2870)
/product="trna-Phe"
complement(2884..4254)
/gene="trnE"
complement(2884..4254)
/gene="trnE"
/codon_start=1
/transl_table=1
/product="trna methyltransferase"
/protein_id="AAC38101.1"
/db_xref="GI:2827009"
/translation="MIRNDITIAQVYCPKRSANGILRVSGIHANQVAFVAVICKPKPR
FATYSKFEFFESKRVVLDGISTLWFPAPSLTGEDVLELOGHGPFIMDLIRLCLAKN
IKIRIAPQGEFCORAFNLNGKIDILQAEALDILNSETESVVRASLNSLHNEFSYLOK
IIRKLEEFRTIASIDFSEENIDPDFNFIIMSPERKLNKPKINIVYSEGLIIEA
KRIVIGPNAKSSLLNLVNCSDRAIVLPTPTDXYENINIGISCEIYDAGL
RETBEDIKERIGIORSMEIKNSDHYLYMDKTSISLDOKKTSIOPKQISSYNIQVY
FVINKDLVDFECGIVTKIENMLTISALTGQCIDILKRLHSRQKDSQEGFIARR
RLHIOIDLSYCELLAKQKMWLKYKNIELLESNLITNKLIGETITGETSSDLKRLFS
TFCIGK"
complement(4297..5907)
/function="unknown"
/transl_table=1
/codon_start=1
/product="60-kDa protein"
/protein_id="AAC38102.1"
/db_xref="GI:2827010"
/translation="MELORNFPIFAFLFVSFELLQAWQSQSLKKKKNEEFSFHLN
HKEDRKQIITKNDVLELVINMGQDIEEASLDYKQLNSPENLKLLEKSPVVOA
QSGILGQDGNSTNKLIRPLYSAKKNFEELGHNEKELRVPTITAKNGITVINFLIK
PGKYNVEYEDINNLSNKKLELNFGLKQVFLPKNRDITSGFALQTRGAAYSS
DAKEYEKFPDNIANHEHLIITTHGWTAMQOYFVAVAVDPDTNISNIIITSLDND
GIATIGYKSSLINIPNSRYIISKSLMIGCEQOEMALVAPNDLTVYDGFLEFLSQP
LFKLLSTIHNFIGNWFSIILTFPMKAITYPYTKAOTSMKRELOPIINELKKNF
GHDQKRSKTEIMALYKREKINPLGGCLPVFIQMPIFSLYMLIGSVELRRAAPFLFWI
KDLSDQDPYVYLPITPKGLTMEFTORTSSNNISDPFOCKIMHMFIFITVFLEPPSGL
VLYIYNSLVITIIQKTYLSNFKKNQK"
complement(5910..6167)
/transl_table=1
/codon_start=1
/product="110-kDa hypothetical protein"
/protein_id="AAC38103.1"
/db_xref="GI:2827011"
/translation="MVKLSTIVPCLTFEVSITYONYISFEMPSRCRYPTCSITYMIS
LRKFGVIGKILITILRLFKCHPLHGGEDLVLKIKDKSEY"
complement(6146..6490)
/gene="trnA"
complement(6146..6490)
/gene="trnA"
/codon_start=1
/transl_table=1
/product="RNase P protein component"
/protein_id="AAC38104.1"
/db_xref="GI:2827012"
/translation="MLNFTFKKSKLLKSTNFQYVFSNPCKNKFPHINILGRSMLGH
PRLGISLRKNIKHAHYRRNKIKRLIRETFRLLOHRLISMDEVIAKKNIYVLNKKIY
NILEYIMSNYOR"
complement(6504..6647)

```
gene
CDS
    /gene="rpmH"
    complement(6504..6647)
    /gene="rpmH"
    /codon_start=1
    /transl_table=11
    /product="50S ribosomal subunit protein L34"
    /protein_id="AAC38105.1"
    /db_xref="GI:2827013"
    /translation="MKRTQPSILKRNHSGFRTMATKNGRIILSRRAKLRTLTIV
    SSK"
    7066..8430
    /gene="dnaA"
    7066..8430
    /gene="dnaA"
    /codon_start=1
    /transl_table=11
    /product="DNA biosynthesis initiating protein"
    /protein_id="AAC38106.1"
    /db_xref="GI:2827014"
    /translation="MSLCIMKQCCLDRLOSELPTSEFMSWIRSLKAKLNNILLEYAPN
    QPILDWYADKYLHFKTLIDPFGSNSPFIKFKYQTSKFKFKKNILOKIQNLNAPK
    IMDKIPLEKSSHSNINKKHSFENFEGKSNOLAASQVAKNPGNSYNPLFLYGG
    TGLCKTHLHAIIGGILAYKYNKIIYHSERFYODAVKALQNNNAIEKFKLYSVD
    LLIDIDIOFPAKERSOEFEFTFNALLEGNOOIIITSDRYPKETNGVEDRLKSRGAC
    LTVADPELEFTRVAIIKIKADENNIVLSDPAFPIAKHLRSNVPELEGALNRYVNA
    NPTHRSTVEFVRALDIIALQEKIYTIANIQTVAEYIKIYKIVADLSRRRSVAR
    PRQAMAMAKELIHNLSLPEIGDAFSGRDTITVLHACRKIEQURKENHDIKEDFSMLIR
    TLSV"
    8435..9535
    /gene="dnaN"
    8435..9535
    /gene="dnaN"
    /codon_start=1
    /transl_table=11
    /product="DNA polymerase III beta subunit"
    /protein_id="AAC38107.1"
    /db_xref="GI:2827015"
    /translation="MKFTIIONDILTKNLKKTIRVLYKNSPFIENILIOVEDGTLSL
    TTYNLEIETLISKIEITRTYIPGKTTISGRKTLINICRTISEKSKITMQLKNNKMTISSE
    NSNTILSTLSADTFPNHONFDYISKFDISSNILEMIEKTEFSWGRQVRYLNGMLL
    EKKKFLKSVATDGYRLAISYQDKDINPFSIIPNKAVMLLKLNTQPOLNII
    GSNRIYTKNLIPTTQLEGEYDPYKSVLFKEKKNPITINSILKSLRLVAIIAHE
    KFCGIEIKENGKFKVLSDNQEEAEDELFEIDYGEKIEISINVIYLDVIYNNKXSE
    NIALFLNKSXSIIQEAENSSNAVYVWLKR"
    9690..12101
    /gene="gyrB"
    9690..12101
    /gene="gyrB"
    /codon_start=1
    /transl_table=11
    /product="DNA gyrase subunit B"
    /protein_id="AAC38108.1"
    /db_xref="GI:2827016"
    /translation="MIDYDSSKIKILRGLDVAVRRKPGMYIGTDDGSGLHHMVEIV
    DNSIDELAGFCFEIKVYIHSDNSYSIKDDRGIPPTDIHPEKISAAEIVIMVYHSGG
    KEDNASKISGGLGVGISVYNALSSEKELIYKKKYKROOYENGNEPSPICIGAT
    EITGTIYRFWPSYKIFTRKTOFOYEILSKRLRELSFNSNTLIHLKDRINDEKTYH
    KGGIKAFKFLKNNKPLINLNIYFRCKNQIELEIALQNNNSHQENILCFTNTPQK
    DGGTHLSGFRAGITRTLMLEIRGYNKKNKTIITGEDTREGTLAISIKIPDFSS
    QTKRLIVSEVSYTESLLENLIEYLENDIAKFILOKTIINSAKIREAARRAEIN
    KKGKGLDGLTLPKGLSDQENDPKSEIYELVSGSAGSAGAKGRNRKNOAILPLKGI
    LNVOKTFDKIILISOELSLTALGCSITRSEYSDIKRYNHIITMPDADGAHIRT
    LLTFYRQDELIRKGGYVYTAOPPLVKKKQKQKQKTKNDEMKKYOIKIALKEIVI
    KNNKNNCNKILKEFOKIVSDFNHIOIKMKKNKNYFPELLNELIYHRLYNLKNENY
    VQNNIEKLEKLNKDKNNNTIYTSKIKENDSDIFEPSIKLSRYANHQYDLDKNFLE
    SKEYFLITNLGEKFKFOENENFEKGEIKYKINDIKNTLEMLIKETKRGFEVORYGK
    LGENPQDLNMTYNNPETRNMLQYTIKDAISANNLFTLMGDIVPRRRKFIQKNALKA
    ENIDI"
    complement(12206..12622)
    /gene="atpC"
    complement(12206..12622)
    /gene="atpC"
    /codon_start=1
```

```
Query Match
Best Local Similarity 1.3%; Score 25; DB 1; Length 34699;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 424 actaaagcgatggaatggttcgtg 448
Db 14881 ACTAAAGCGATGGAATGCTCTG 14857
```

Search completed: October 28, 2001, 17:56:17
Job time: 8411 sec

Query Match	Best Local Similarity	Score 1972:	DB 21:	Length 1972:
Matches 1972: Conservative	100.0%;	Pred. No. 0;	Mismatches 0;	Indels 0; Gaps 0
0Y	1	accggcgccattggtcagtcagtcgcttcgcatcttcatgatgacacgtaagaacattatc	60	
Db	1	agcgggccattggtcagtcagtcgcttcgcatcttcatgatgacacgtaagaacattatc	60	
0Y	61	acatggtcaaaaagaactgaactgaattatgaacaaacacatattctcggtatctgtagc	120	
Db	61	acatggtcaaaaagaactgaactgaattatgaacaaacacatattctcggtatctgtagc	120	
0Y	121	acaacaagcattagtggttatttcgcagtagagtttgcgttacttagaagaagtgcattaga	180	
Db	121	acaacaagcattagtggttatttcgcagtagagtttgcgttacttagaagaagtgcattaga	180	
0Y	181	tcgtattggttcatttgaatcagcacttttagagtagtgtaaaccaataactatgtatatt	240	
Db	181	tcgtattggttcatttgaatcagcacttttagagtagtgtaaaccaataactatgtatatt	240	
0Y	241	tatggtgtagttaaccacactctgcacattatcacatgatagaatgaagaagtcattaaagg	300	
Db	241	tatggtgtagttaaccacactctgcacattatcacatgatagaatgaagaagtcattaaagg	300	
0Y	301	catcttgtagcttccaagaacacacagtcgctgtagtaagttaacactttaatgtagagac	360	
Db	301	catcttgtagcttccaagaacacacagtcgctgtagtaagttaacactttaatgtagagac	360	
0Y	361	aaaatggcgggtgctcaaaagataaagaacaaatgcgagctgttaaagaatcacaaaaa	420	

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 28, 2001, 15:34:11 ; Search time 1454.83 Seconds
(without alignments)
12813.200 Million cell updates/sec

Title: US-09-545-199c-3

Perfect score: 1972

Sequence: 1 agcgggcattgctcagc.....gtcttattcttgatgataa 1972

Scoring table: OLIGO_NUC

Gapop 60.0 , Gapext 60.0

Searched: 10228115 segs, 4726426750 residues

Word size : 25

Total number of hits satisfying chosen parameters: 0

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database :

EST:*

1:	gb_est1:*
2:	gb_est2:*
3:	gb_est3:*
4:	gb_est4:*
5:	gb_est5:*
6:	gb_est6:*
7:	gb_est7:*
8:	gb_est8:*
9:	gb_est9:*
10:	gb_est10:*
11:	gb_est11:*
12:	gb_est12:*
13:	gb_est13:*
14:	gb_est14:*
15:	gb_est15:*
16:	gb_est16:*
17:	gb_est17:*
18:	gb_est18:*
19:	gb_est19:*
20:	gb_est20:*
21:	gb_est21:*
22:	gb_est22:*
23:	gb_est23:*
24:	gb_est24:*
25:	gb_est25:*
26:	gb_est26:*
27:	gb_est27:*
28:	gb_est28:*
29:	gb_est29:*
30:	gb_est30:*
31:	gb_est31:*
32:	gb_est32:*
33:	gb_est33:*
34:	gb_est34:*
35:	gb_est35:*
36:	gb_est36:*
37:	gb_est37:*
38:	gb_est38:*
39:	gb_est39:*
40:	gb_est40:*
41:	gb_est41:*
42:	gb_est42:*
43:	gb_est43:*

44:	em_esthum10:*
45:	em_esthum11:*
46:	em_esthum12:*
47:	em_esthum13:*
48:	em_esthum14:*
49:	em_esthum15:*
50:	em_esthum16:*
51:	em_esthum17:*
52:	em_esthum18:*
53:	em_esthum19:*
54:	em_esthum20:*
55:	em_esthum21:*
56:	em_esthum22:*
57:	em_esthum23:*
58:	em_esthum24:*
59:	em_esthum25:*
60:	em_esthum26:*
61:	em_esthum27:*
62:	em_esthum28:*
63:	em_estin1:*
64:	em_estin2:*
65:	em_estin3:*
66:	em_estin4:*
67:	em_estin5:*
68:	em_estom1:*
69:	em_estom2:*
70:	em_estov1:*
71:	em_estov2:*
72:	em_estp11:*
73:	em_estp12:*
74:	em_estp13:*
75:	em_estp14:*
76:	em_estp15:*
77:	em_estp16:*
78:	em_estp17:*
79:	em_estp18:*
80:	em_estp19:*
81:	em_estp110:*
82:	em_estro1:*
83:	em_estro2:*
84:	em_estro3:*
85:	em_estro4:*
86:	em_estro5:*
87:	em_estro6:*
88:	em_estro7:*
89:	em_estro8:*
90:	em_estro9:*
91:	em_estro10:*
92:	em_estro11:*
93:	em_estro12:*
94:	em_estro13:*
95:	em_estro14:*
96:	em_estro15:*
97:	em_estro16:*
98:	em_estro17:*
99:	em_estro18:*
100:	em_estro19:*
101:	em_estro20:*
102:	gb_est25:*
103:	gb_est26:*
104:	gb_est27:*
105:	gb_est28:*
106:	gb_est29:*
107:	gb_est30:*
108:	gb_est31:*
109:	gb_est32:*
110:	gb_est41:*
111:	gb_est42:*
112:	gb_est43:*
113:	gb_est44:*
114:	gb_est45:*
115:	gb_est46:*
116:	gb_est47:*

```

117: gb_est48:*
118: gb_est49:*
119: gb_est50:*
120: gb_est51:*
121: gb_est52:*
122: gb_est53:*
123: gb_est54:*
124: gb_est55:*
125: gb_est56:*
126: gb_est57:*
127: gb_est58:*
128: gb_est59:*
129: gb_est60:*
130: gb_est61:*
131: gb_est62:*
132: gb_est63:*
133: gb_est64:*
134: gb_est65:*
135: gb_est66:*
136: gb_est67:*
137: gb_est68:*
138: gb_est69:*
139: gb_est70:*
140: gb_est71:*
141: gb_est72:*
142: gb_est73:*
143: gb_est74:*
144: gb_est75:*
145: gb_est76:*
146: gb_est77:*
147: gb_est78:*
148: gb_est79:*
149: gb_est80:*
150: gb_est81:*
151: gb_est82:*
152: gb_est83:*
153: gb_est84:*
154: gb_est85:*
155: gb_est86:*
156: gb_est87:*
157: gb_est88:*
158: gb_est89:*
159: gb_est90:*
160: gb_est91:*
161: gb_est92:*
162: gb_est93:*
163: gb_est94:*
164: gb_est95:*
165: gb_est96:*
166: gb_est97:*
167: gb_est98:*
168: gb_est99:*
169: gb_est100:*
170: gb_est101:*
171: gb_est102:*
172: gb_est103:*
173: gb_est104:*
174: gb_est105:*
175: gb_est106:*
176: gb_est107:*
177: gb_est108:*
178: gb_est109:*
179: gb_est110:*
180: gb_est111:*
181: gb_est112:*
182: gb_est113:*
183: gb_est114:*
184: gb_est115:*
185: gb_est116:*
186: gb_est117:*
187: gb_est118:*
188: gb_est119:*
189: gb_est120:*

```

```

190: gb_est110:*
191: gb_est111:*
192: gb_est112:*
193: gb_est113:*
194: gb_est114:*
195: gb_est115:*
196: gb_est116:*
197: gb_est117:*
198: gb_est118:*
199: gb_est119:*
200: gb_est120:*
201: gb_est121:*
202: gb_est122:*
203: gb_est123:*
204: gb_est124:*
205: gb_est125:*
206: gb_est126:*
207: gb_est127:*
208: gb_est128:*
209: gb_est129:*
210: gb_est130:*
211: gb_est131:*
212: gb_est132:*
213: gb_est133:*
214: gb_est134:*
215: gb_est135:*
216: gb_est136:*
217: gb_est137:*
218: gb_est138:*
219: gb_est139:*
220: gb_est140:*
221: gb_est141:*
222: gb_est142:*
223: gb_est143:*
224: gb_est144:*
225: gb_est145:*
226: gb_est146:*
227: gb_est147:*
228: gb_est148:*
229: gb_est149:*
230: gb_est150:*
231: gb_est151:*
232: gb_est152:*
233: gb_est153:*
234: gb_est154:*
235: gb_est155:*
236: gb_est156:*
237: gb_est157:*
238: gb_est158:*
239: gb_est159:*
240: gb_est160:*
241: gb_est161:*
242: gb_est162:*
243: gb_est163:*
244: gb_est164:*
245: gb_est165:*
246: gb_est166:*
247: gb_est167:*
248: gb_est168:*
249: gb_est169:*
250: gb_est170:*
251: gb_est171:*
252: gb_est172:*
253: gb_est173:*
254: gb_est174:*
255: gb_est175:*
256: gb_est176:*
257: gb_est177:*
258: gb_est178:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result	Query	Score	Match Length	ID	Description
No.					

No matches found

Search completed: October 28, 2001, 17:10:56
Job time: 5805 sec

THIS PAGE BLANK (USPTO)

Query Match	63.7%;	Score 1255.6;	DB 3;	Length 13820;
Best Local Similarity	78.3%;	Pred. No. 1.1e-235;		
Matches 1536;	Conservative	0;	Mismatches 414;	Indels 12;
				Gaps 2

QY	11	ttggccagcttcggttcggtatctctatgatgcacgtaagaacatatacaatgtgtccaa	70
Db	9952	TTTGGCAGCTTTGGCATCGAATCTGAGATGCTCGACTCGTAGCAACATTTCTCCACGGTGA	9893
QY	71	aaagtaactcgaatttatbgaacaaacaaactactctccattactctbtaagacaacaaga	130
Db	9892	AAAGTAACTCGAATTTATTTAAACAAAACAAATATCCAGCCATTATCTGTGACAGAACAGG	9833
QY	131	ttagtgctattctgcagtagagcttctgttacttagaagaagtggaactagatcgtattgt	190
Db	9832	GTTATATTTGTTTGCTGTTGAATTTTGGTATTTGGATGACGCGAATATTCAAAATTTGCA	9773
QY	191	tcaattgaactcagcaactcttagagatgatctacatcaactatgtagtttatcgtgtg	250
Db	9772	AGTTTGGAAACCGCACCTTTTAATTTATTTCTAACCGTATATCCGCTGATTTATTTAGCAAGAG	9713
QY	251	ttaacccaactctggcgcaattacatgatbgaattaaagaagtcataaagaagcattctggat	310
Db	9712	CTTAAATTAACCCGCTATCATATATGATGAATCAAAAGTACTTTTAAAAAGATTTTAGAT	9653
QY	311	agcttcaagaacaaacagtcgctgtgtaagttaacacttaaatbgaagacaatatgycag	370
Db	9652	GGTTTTTAAAGCGCATGTCTCTTGTA-----GTAAAGGAGATTAAGATGGCAG	9604
QY	371	gtcttaagaagataagaaccaaatccgcgagtgtaaaagatcaccaaaatattactaaag	430
Db	9603	GTGCAAAAGAGTAATAAACCAAAATTTGCCAGTGTCAAATGACAAAATAATCCATAGG	9544
QY	431	cgaibgaatbgtctgtccttcgaaaaatgcytaaaacgcgaagcagcttctcttcac	490
Db	9543	CAATGCAAAATGGTGGCAACCTCGAAATTCGTAACCAAGATTCGATGGCTGCAATCTC	9484
QY	491	ggccttatcaagaacaatcagctaaagtgatagcagaagcttccaagaacagatctgtgt	550
Db	9483	GTCCGATTTCTGAACATATCCGTAACGTTATTATGTCATGTGTCATAGGCAGATGCGTT	9424
QY	551	acaagcaccattttagtgcgacgcgaagtaaaaaagtggcatagtattgtgtcca	610
Db	9423	ATAAACCTCGCTTCTAGTGWAGCGCGAAGTCAAGAAATCGGATCTTGTTATTTTCAA	9364
QY	611	cagctgtgctcttcttggttggtccttaagaatgaactgtttaaactgtataatgtaa	670
Db	9363	CAGATCTGTGGATGTGTGGTGTTAAATGTTAAATTTATTTATTCAAAACCACTTAAACCAA	9304
QY	671	tgaagaatbgaagaagaagaatgcttcggtccaatbgaattaaatcoggttctaatacta	730
Db	9303	TAAAAAATTTGGAAAGAAACAAATATTTCTACAGATTTTGGGCTTATTAAGTTCAAAAGGA	9244
QY	731	tcaactttccaactcttgggaataaattttaacccaagaatcagatattgtgata	790
Db	9243	TTAGTTTTTCCGTTCTTTGGATTAAATATTAATAAAGTCAAGTTCGTTCTGTGTAAGCGATA	9184
QY	791	ctccctctgttcagcagttaatctgtctcagtcgaatctctatgatbgaatgactataaaaag	850
Db	9183	CGCCCGCTCTGAAGATTTATTTGGTGGCAAAATTCATATGTTTGATCTTATGCTATATG	9124
QY	851	gggaagatagatgttgtagttattagttataaacaattttaacacggttcgcgaaaagc	910
Db	9123	GTGAAATTTGATGCGTTTATTTATTTGCAATACATATAATTTGTTAATTCGATGTCCAAAAGC	9064
QY	911	cagatctbgaanaaattaatccattcaacgaatctagatattgatbgaattagcgcgaaga	970
Db	9063	CTGTTGTACAACTATTTAGTTCTTTTACCAAGAAATCTTAAGACATCATTTTAAATGCAAGAC	9004
QY	971	aacaagtttggatcataattcgaacccgtagtgcgaagaatattatagataattatggt	1030
Db	9003	AACGACTTTGGATTTATCTTTATATGAGCCAAACCAAAATGCTATTTAGTATGACCTTTTAA	8944

Accession	Gene	Position	Strand	Sequence	Length
1031	ttgcta	1031	+	ttgcta	1090
8943	tttccgtt	8943	+	tttccgtt	8884
1091	ccgctg	1091	+	ccgctg	1150
8883	cgctgc	8883	+	cgctgc	8824
1151	tacag	1151	+	tacag	1210
8823	tggcg	8823	+	tggcg	8764
1211	ttgcg	1211	+	ttgcg	1270
8763	tacgc	8763	+	tacgc	8705
1271	acaat	1271	+	acaat	1330
8704	acaa	8704	+	acaa	8645
1331	tga	1331	+	tga	1380
8644	cgat	8644	+	cgat	8585
1391	tga	1391	+	tga	1450
8584	agta	8584	+	agta	8525
1451	aaat	1451	+	aaat	1510
8524	aaac	8524	+	aaac	8465
1511	cgat	1511	+	cgat	1570
8464	tgt	8464	+	tgt	8405
1571	tcac	1571	+	tcac	1630
8404	ccac	8404	+	ccac	8345
1631	aata	1631	+	aata	1690
8344	ttt	8344	+	ttt	8285
1691	tgt	1691	+	tgt	1756
8284	cgt	8284	+	cgt	8225
1751	gaac	1751	+	gaac	1810
8224	gcatt	8224	+	gcatt	8165
1811	ctata	1811	+	ctata	1870
8164	ctata	8164	+	ctata	8105
1871	gaac	1871	+	gaac	1930
8104	gaac	8104	+	gaac	8045
1931	attc	1931	+	attc	1972
8044	attc	8044	+	attc	8003

Source	Organism	Reference Authors Title	Journal	Reference Authors Title	Journal
source	Pasteurella multocida.	Fuller, T.E., Kennedy, M.J. and Lowery, D.E.	Unpublished	Fuller, T.E., Kennedy, M.J. and Lowery, D.E.	Unpublished
gene	Pasteurella multocida.	Identification of Pasteurella multocida virulence genes in a septicemic mouse model using signature-tagged mutagenesis	2 (bases 1 to 870)	Submitted 24 FEB 2000, J.D. Discovery Research, Pharmacia & Upjohn Animal Health, 7925-25-434, 7000 Portage Road, Kalamazoo, MI 49001-0199, USA	2 (bases 1 to 870)
CDS	location/Qualifiers				
	1. .870	/organism="Pasteurella multocida"			
	/db_xref="taxon:747"				
	/gene="atpC"				
	1. .870				
	/gene="atpG"				
	/codon_start=-1				
	/translation=-1				
	/product="ATP synthase F1 gamma chain"				
	/protein_id="AA68408.1"				
	/db_xref="GI:7716505"				
	/translation="MACAKEIRFKIASVKSNOITKAMENYVAAKMKRTOBRMSSRPVSEIRIVSHVSKATGTYKRPHPVDEKVGKVIYVSPRGCLGLNVLFRKVLNE MKERKEDVSVOLSLIOKSIINPFOSLGILITDDSGIDTPSEOLIGSVNSIDAN KKEEDVYVLVYNKFIWMSOKPVLELIPPELDNDELGRKQVMDIYEPDKVLL DNLEVRLESQVYQAAVENLASEQAAVMVAMKATDVAAGNLINELQVLYNARQASITIT NELNELYVAGAAAT"				
BASE COUNT	305 a	132 c	177 g	256 t	
ORIGIN					
Query Match	44.1%; Score 870; DB 2; Length 870;				
Best Local Similarity	100.0%; Pred. No. 2.8e-160;				
Matches 870; Conservative	0; Mismatches 0; Indels 0; Gaps 0;				
Y	364 atggcaggtgctaaagataagaacccaacaaatcgcgaggtgtaaaagtacacaaaatt	423			
Db	1 ATGCGAGGTGCTAAAGAGATAGAACCAAAATCGCGAGTGTAAAGATACAAAAAT	60			
Y	424 actaaagcgatgaatagt	483			
Db	61 ACTAAAGCGATGAATAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	120			
Y	484 tcttcacgaccttlttcgaacaatactagctagctagctagctagctagctagctagct	543			
Db	121 TCTTCACGCCCTTATTCGAACAACATACGTACGATTAACCGACGTTTCCAAACACG	180			
Y	544 atcgtgttaacgaatccaattttagtgcgtgcgaagtaaaaaaagtggcgatgtgt	603			
Db	181 ATTGGTTACAGCATCCATTTTGTAGTGATCGCGAAGTAAAAAAAGTGCGCATGTGT	240			
Y	604 gtgcgcacagtcgt	663			
Db	241 GTGTCCACAGATCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	300			
Y	664 aatgaatgaagaatggaagaagaagaatgttccgttcaattgagtttaatcggttct	723			
Db	301 AATGAATGAAGAATGGAAGAAGAAGAATGTTCCTTCATTCAGTTAATCGGTTCT	360			
Y	724 aaatctatcaacttttccaactttgggaattaaaatttaacccaagatcaggtat	783			
Db	361 AAATCTATCAACTTTTCCATCTTTGGGAATTTAAATTTTAAACCCAGATTCAGGTAT	420			
Y	784 ggttaactccctgtgtgacagtttaattgtgttcagtcacattctatgtatgattat	843			
Db	421 GGTATACCTCCCTGTGTGACAGTTAATGTGTGTGTGTGTGTGTGTGTGTGTGT	480			

```

OY      844   aaaaagggaagttagctgttgattagttataacaaatattacaacgactgcg     903
DB      461   AAAAAAGGGGAAGTAGTGTGATTATTTAGTTTTTAACAATTTATTAAACGCATGTCG    540
OY      904   caaaagccagtatltggaaaaattaatcccatcacagaattagataatgatgaattagc     963
DB      541   CAAAAGCCAGTAGTGGAATAAATTAAATTCATACCAGAATTAGATAATGATGAATTAAGC      600
OY      964   gaagaaacaagaatltgggatatatttcaagaacctgatatcgaaagtattatagataat     1023
DB      601   GAAAGAAAAACAAGTTGGGATATTAATTTAACGAACCTGATGCCAAAAGTATTAATTAATAT       660
OY      1024  ttattggttcgtatttagaacatccaagtttaacagaacgagtggaaaacctgctct     1083
DB      661  TTATTGTTGTTGTTATTATTTAGCATCTAGGTTTATCAAGCACGAGTTGAANAACCTTGCTTCT     720
OY      1084  gagcaagccgcctcgaaatlygttcgccatlgaaagagaagaacagataacgcagtaactaat     1143
DB      721  GAGCAAAGCGCGCTCGAATGGTGCCCATGAAGAAGCAGCAACGATPAACGAGTAACTTAAT       780
OY      1144  aatgagttacagttatgctctaacaagaactcgctcaagaagttattaaagaatlaaat     1203
DB      781  AATGAGTTACACTTGTGCTATPMAACAACCTCGTCAACGACAGTATTACAAATGAATTAAT       840
OY      1204  gaaattgtttccggtgctgcagcagaactttaa     1233
DB      841  GAATGTTGTCGCGTGACACACCAATTAA        870

RESULT      4
LOCUS              ECOHAMP           7141 bp          DNA            BCF             20-DEC-1995
DEFINITION      E.coli H+ ATPase alpha, beta, gamma, delta and epsilon, and
                integral membrane proton channel at D-man6C subunit genes,
                complete cds.
ACCESSION      M25464
VERSION        M25464.1 GI:146318
KEYWORDS       H+ ATPase; integral membrane proton channel.
SOURCE         Escherichia coli
ORGANISM       Escherichia coli
REFERENCE      1 (bases 1 to 7141)
AUTHORS       Kanazawa,H. and Futai,M.
TITLE         Structure and function of H+-ATPase
               Escherichia coli H+-ATPase
               Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
               Escherichia.
JOURNAL       Ann.N.Y.Acad.Sci. 402:45-64.(1982).
MEDLINE       83176724
FEATURES       location/qualifiers
                 1..7141
                 /organism="Escherichia coli"
                 /db_xref="taxon:562"
                 165..557
                 /note="14 Kd ORF; putative"
                 /codon_start=1
                 /transl_table=11
                 /product="unknown protein"
                 /protein_id="AA83868.1"
                 /db_xref="GI:1128958"
                 /translation="MKNVGSYSLSRVNARKLILVQLLVIASGLFSLKDPFWASVA
ISGLAVELPLVLEMFAMRHOATPKAGRVMFAFGGEAFKVLIVLLVALAVIK
AVEFLVTWTVLVVOILLAPVINNG"
                 566..1381
                 /codon_start=1
                 /transl_table=11
                 /product="integral membrane proton channel F0 subunit A"
                 /protein_id="AA83869.1"
                 /db_xref="GI:146319"
                 /translation="MASENNTPPOVIGHHNLNIOLDLFESINDPONMPATFPWTIND
SMPSVIYLGLVLPFRSVAKKAISGPGRFORIAIELYGFVNSGYADTHGSKLIA
PLATITVVWFLNMIDLPILDIPYIAEHLVGPARVPSPADVNTLSMALGVFTLL
IFLATIMKGIGTFTEKRSTFNHNAFIPIVNLIIEGVSIISKVSIGLRIFGMVAG
```

CDs

ELIFLLIAGLLPMSQMLNVPMAIFHILLITTLQAFIMVLITVLSMASEEN"
1428..1667
/codon_start=1
/transl_table=11
/product="Integral membrane proton channel F0 subunit C"
/protein_id="AA83870.1"
/db_xref="GI:146320"
/translation="MENLMDLLYMAAYMMGLAICATIGITLGGKFLFEGAAKRPD
LIPILRQFPIVGMGLVADIPMAIVAGLGLVYFAVA"
1729..2199
/codon_start=1
/transl_table=11
/product="Integral membrane proton channel F0 subunit B"
/protein_id="AA83871.1"
/db_xref="GI:146321"
/translation="MNLNATLIGQALAFVFLFCEMKYVYPMPLMAIEKROKEIADGL
ASAEHAKDLDLAKSAFDOLKKAKEKOVITIEQANRRSQIIDENAEAEKERTKIV
AQAQAEIEERARRAREELRKQVALLAVAGAEKIIERSVEDEANSNDIVDKLVAEL"
2214..2747
/codon_start=1
/transl_table=11
/product="H+ ATPase F0 delta subunit"
/protein_id="AA83872.1"
/db_xref="GI:146322"
/translation="HSEFTVAPYAKAAPFAVEHOSYVERMDLAFAAEYTKNEOM
AELLSGALAPETLASFIAVCGEQLDENQGLIRYMAENKRLNLPVLIQFTHIRAV
SEATLEVADYISAAALSEQOLAKISAMERKRYKLNCIDKSVMAGVITIRAGDAVY
DGSVRGRLERLADYQGS"
2760..4301
/codon_start=1
/transl_table=11
/product="H+ ATPase F1 alpha subunit"
/protein_id="AA83873.1"
/db_xref="GI:146323"
/translation="MQLNSTESELKORIAQFNVSSEAHNEGTIVSDGVRINHGL
ACQCEMSLISGNRYATAINLEBDVSAGVAMPADLAEGMKICIGRIILEYVGGH
LGRVNTIGAPIDCKGPLDHDGSESAVLAIPGYEORSVQDPQVOTGKAVDSNPIG
RQGEELIIDCROTKTALADIKITMDRSGIKCTIVAIQCASTISNVYKLEHGHAL
ANTIVVAASASALQYLARPMVALMEGYREDREGLIYDLSKOAVARQISLL
LRPREFREPNDIVFLHSLRLEERAARVAEVEAFPTGKGVKGTSLTALPIETQAA
GVDSAPFENISITDGOIFLETENFNAGIRPAVNPGISVSGAGCAQTLMKLSGGG
IRFALAOVEELIAESQFASDLDDATRKQDLHOKYTELKOKOYAPMSVAQSIYLFRA
AERGVIADVELSKIGSEFAALLAYVDROHAPLMQDINGTGVYDEBCKLGLDSPK
ATQSW"
4352..5215
/codon_start=1
/transl_table=11
/product="H+ ATPase F1 gamma subunit"
/protein_id="AA83874.1"
/db_xref="GI:146324"
/translation="MAGAEIRSKIASVONTOKITKAMENVAASMKRKSQDMAASRP
YATMKRKYVIGHLAHGNIKEKHPYLEDRDQKRVAGYLVSTDRQLCGGLINLPEKLLMA
KMTWTKQOCLDAMIGSKGVSEFNSVDGNAAGYVAGMDNPSSELLGPEKVIQMAAA
DEBRLDKIVTSYKSPFINMSOYPTISOLLPLPASDDDLKRSMPYIYEPPKALDLPL
ILRRYBSQYIGCVENLASBOAARVAKKATDNGSGSLIKELQLYVKKAAQASITQEE
LREIVSGAAV"
5242..6624
/codon_start=1
/transl_table=11
/product="H+ ATPase F1 beta subunit"
/protein_id="AA83875.1"
/db_xref="GI:146325"
/translation="MATGRIVQYIGAVVDVEFPDAPVRYDALEVONGNRELVLENG
QOGLGIVRTIAMSDDGLRGLDVKDLKYLEVPGKATLGRINNVLGEPEVDMKGELE
GEERRAIIHRAAPSAEYELNSOILETGIKVIDLPMKAGKGVGLFGAGVKGTVNNR
MELIRNIAIHSGISYFAGYGEERTRGNDFYHEMDSNVYIDKVSLYTQAMNPPGNRI
RVALATITAEKFRDEGRDVLFDVNIYTTLAGEVSAALLRMPSAQIYOPTLAIEG
GVLOERTSTKRTGISYQAVVYPAADLDLDPATFALDFTVALSQIASLGIYTPK
VDPLDSRQDLPLVGOEHYDPAKGVOSIILROYELKQIILAILEMDELSBDDKLVVPA
BARTIOEISOPFVAEVLFGSPGKRYVSLKTIIRGKGLIMEGEYDHLPEQAFYVWSI
BEAVERKAKKL"
6645..7064
/codon_start=1

CDs

CDs

CDs

CDs

[illegible]

Db 4815 TGATGTTGAGCGCTGACGAGGAGCGCTGTGAGCAACGTTTACATGTCAGCAACAAT 4874
 QY 887 ttattacaacgagtgtcgcaaaagccagtlattggaanaaataatcattaccaggaattag 946
 Db 4875 TTAATTAACACCATGTCTCGAGTTCACACCATTCAGCCAGCTGCTCCGTTACCCGCAATCA 4934
 QY 947 ataataatgaattaggcgaagaagaagaattgggagttatatatttcgaacacgtatgcga 1006
 Db 4935 ATGATGATATATCT-----GAACATTAATCCTGGATTAACCTGTACGTAACCCGATCCGA 4988
 QY 1007 aagttattatagaataattatttgctgtattatagaatcctcaggttatacgaacgacaa 1066
 Db 4989 AGCGCGTGTGATACCGTCCGTCGTATATGCAATCTCAGGTTTATCAGGGGTG 5048
 QY 1067 ttgaaacacttcttcttgacgaacgcctgcgaatggctgcacatgaagaagcaagcaagata 1126
 Db 5049 TTGAAATACTGGCGACGACGACGCGCGCTATGCTGGGATGAAGAAAGCCGACCCAGCA 5108
 QY 1127 acgcaggttaacttaattatgaattagttacagttatgatatacaagaactcgtcaagcaagta 1186
 Db 5109 ATGCGCGGACCTGATTTAAAGACCTCAGTTGATTAACAACAAAGCTCGTACAGCCAGCA 5168
 QY 1187 ttacaataatgaatnaatgaatgttgcgcgtgcagcagaacttaacaa-----a 1237
 Db 5169 TTAACCTAGAGAACTCACCGAGATGCTCGGGGGCGCGCGGTTTAAACAGGTTATTTGCG 5228
 QY 1238 tagaagatcgtgtaaatgacacgtggaanaattgtacaatacattcgtgtgcgttatagcgt 1297
 Db 5229 TAGAGATTTAAGATGCTACCTGGAAGATTTGTCAGGTAATCGCGCGTAACTGATACGT 5288
 QY 1298 tgaattcccaacaagaatgacagtaacaaagatatatagtcctt-----aaatttga 1348
 Db 5289 CGAATTCCTCCAGAGATGCGGTACCGCGCGGTACGATGCTCTGAGCTGCAAAATGCTAA 5348
 QY 1349 aacaaggtttagtacttgaagttcaacaacaataatggtgtgtgtgtgtgtgtgtgtgtgt 1408
 Db 5349 TGAGCGCTGTGCTGTGGAAGTTTCAGCAGCAGCTCGGGGGGATGCTAGTACCTACCTCG 5408
 QY 1409 aatggatcatctgtatgtatgaacgcggtttaaagcgttaacaatacgaataacccaat 1468
 Db 5409 AATGGGTTCTCTCGACGCTGCTGCTGCTGATGTAAGAACCTCGAAMACCCGAT 5468
 QY 1469 tctgttccagtggggaagcaaaaacattgggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1528
 Db 5469 TGAAGTCCCGGTAAGTAACGACTGTGGCGGTATCATGATGATGATGATGATGATGATGAT 5528
 QY 1529 cgatgacgaaggtgaatcgtgtgcagaagaagaattgtctatccacgtgtgcgcacaaag 1588
 Db 5529 CGCATGGAAGGCGATCGGTGAAGAGAGGTTGGCGATTACCGCGCACACCTTC 5588
 QY 1589 ttatgaagaacaactcaactgaacttttagaagaacggaataaagtatgcactt 1648
 Db 5589 CTACGGAAGAGCTGTCAAACTCTCAGGAACCTGCGTAACCCGATCAAGTATGACACCT 5648
 QY 1649 agttgttcggttgcgaagaagggtgaagaatgattatcgtgtgtgtgtgtgtgtgtgtgt 1708
 Db 5649 GATGTCCTCGTGTGTAAGGCGGTAAGTGTGCTGCTGCGGTGCGGTGTGTAAGTGA 5708
 QY 1709 aacgctcaatagatgaatataatccgtacacatcgcaatgcaactcaggttactcgt 1768
 Db 5709 AACCGTAACATGATGAGACTCATTCGTAAACATCGGATCGAGCACTCCGGTTACTCTGT 5768
 QY 1769 cttgcggggtagttagtgcgtacgcgtgaaggttaacgaacttcatcagatgaaga 1828
 Db 5769 GTTGTGGGCGGTAGGTAAGTACTGCTGAGGGTAAAGACTTCTACCAAGAAATGACCGCA 5828
 QY 1829 ctctaacgatatagataaaggtgtctcttcttattatggtcaaatgaagaagcaccaggtta 1888
 Db 5829 GTTCAACGTTATCGACAAATATCCCTGCTGTATGCGCAGATGAACGACCCCGCGGAAA 5888
 QY 1889 ccgttaactgtgtgcattacagagcttaactatgctgcggaanaattcgtgtatgaaggtgc 1948
 Db 5889 CCGTCTGCGCGCTTCTCTGACCGGCTGTGACATGCTGTGAGAAATTCCTGTGACGAAAGTGC 5948

QY 1949 tgatgtctattcttcgttgatata 1972
 Db 5949 TGAGCTTCTGCTGCTGTGACAA 5972

RESULT 5
 AE000450/c
 LOCUS
 DEFINITION
 Escherichia coli K12 MG1655 section 340 of 400 of the complete genome.
 ACCESSION
 VERSION
 AE000450.1 GI:1790166
 KEYWORDS
 SOURCE
 ORGANISM
 Escherichia coli K12.
 Escherichia coli K12.
 Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae; Escherichia.

REFERENCE
 AUTHORS
 1 (bases 1 to 11414)
 Blattner, F.R., Plunkett, G. III, Bloch, C.A., Perna, N.T., Burland, V., Riley, M., Collado-Vides, J., Glasner, J.D., Rode, C.K., Mayhew, G.F., Gregor, J., Davis, N.W., Kirkpatrick, H.A., Goeden, M.A., Rose, D.J., Mau, B. and Shao, Y.
 The complete genome sequence of Escherichia coli K-12
 Science 277 (5331), 1453-1474 (1997)

JOURNAL
 MEDLINE
 PUBMED
 9742617
 2 (bases 1 to 11414)
 REFERENCE
 AUTHORS
 Blattner, F.R.
 Direct Submission
 Submitted (16-JAN-1997) Guy Plunkett III, Laboratory of Genetics, University of Wisconsin, 445 Henry Mall, Madison, WI 53706, USA.
 Email: ecollegenetics.wisc.edu Phone: 608-262-2534 Fax: 608-263-7459

REFERENCE
 AUTHORS
 3 (bases 1 to 11414)
 REFERENCE
 AUTHORS
 Blattner, F.R.
 Direct Submission
 Submitted (02-SEP-1997) Guy Plunkett III, Laboratory of Genetics, University of Wisconsin, 445 Henry Mall, Madison, WI 53706, USA.
 Email: ecollegenetics.wisc.edu Phone: 608-262-2534 Fax: 608-263-7459

JOURNAL
 MEDLINE
 PUBMED
 9742617
 4 (bases 1 to 11414)
 REFERENCE
 AUTHORS
 Plunkett, G. III.
 Direct Submission
 Submitted (13-OCT-1998) Laboratory of Genetics, University of Wisconsin, 445 Henry Mall, Madison, WI 53706, USA
 This sequence was determined by the E. coli Genome Project at the University of Wisconsin-Madison (Frederick R. Blattner, director). Supported by NIH grants HG00301 and HG01428 (from the Human Genome Project and NCHGR). The entire sequence was independently determined from E. coli K12 strain MG1655. Predicted open reading frames were determined using Genemark software, kindly supplied by Mark Borodovsky, Georgia Institute of Technology, Atlanta, GA, 30332 (e-mail: markborov@gatech.edu). Open reading frames that have been correlated with genetic loci are being annotated with CG Site Nos., unique ID nos. for the genes in the E. coli Genetic Stock Center (CGSC) database at Yale University, kindly supplied by Mary Berlyon. A public version of the database is accessible (http://cgsc.biology.yale.edu). Annotation of the genome is an ongoing task whose goal is to make the genome sequence more useful by correlating it with other data. Comments to the authors are appreciated. Updated information will be available at the E. coli Genome Project's World Wide Web site (http://www.genetics.wisc.edu). ** The E. coli K12 sequence and its annotations are periodically updated; this is version M54. No sequence changes. Annotation updates: updated gene identifications and products; all new functional assignments courtesy of Monica Riley; added promoters, protein binding sites, and repeated sequences described in reference 1. The unique numeric identifiers beginning with a lowercase 'b' assigned to each gene (protein- or RNA-encoding) are now designated as gene synonyms instead of labels. This should allow them to be searched for in Entrez as gene names.

D	7035	AAAGGACCCGAACCTGCGTAAGCAACGAACAACTATAGCCCGCATGTGCGTTCGACGACACTCT	6978
O	131	ttatgttcatcttcgtagtaggttttggtttaacttagaagaatgtagcttagatcgttatgtt	190
D	6975	CTGGTTCGTTCGCGACGACGAACCGTGGTTACCTGCGCGATGTTGAACCTGCGAAATTGCG	6916
O	191	tcatttaatcagcacttttagagtagtgcataaccatacctagtcgatcttattatcgtag	250
D	6915	AGCTTCGAAGCGCTGCTGCGCTTACGTCGACCGGATCGACGCGCTCGTTGATGCGAAG	6856
O	251	ttaacccaatctgtgcaattcaatagtgaataaagagtagtaattaaagcaattttg	310
D	6855	ATCAACACGACGACCGGTGGCTCAACGACGAAATGGAAGGCAACTGAAGGCACTCTCGAT	6796
O	311	agcttcaagaacaaacagtagtgcgtta-----agttaacact	346
D	6795	TCCATTCAAGCAACCCATCTCGTTAAGCTTGCGCGCTTACGGGCAAGGCCGCG	6736
O	347	ttaatgtagagagaacaaatgycagtggtctaaagagataagaacccaatcgcgagttaa	406
D	6735	GCATTGAGGAGACACTTCATGTCGGCGGCGCAAAAGATAGTGTGTAAGATCCGACCGTCC	6676
O	407	aaagtacaacaaaattaccctaaagcgatgaaatgtgtgtcccggaatctgcgttaa	466
D	6675	AGAACACGCAAAAGATACATAAAGCCATGAGATGTTCCGCCCTTCAAAATGGCGTAAT	6616
O	467	cgcagaagcagatgtcttcttcacgcccattatcgaagaacaaatacgtatgtagcc	526
D	6615	CGCAGATCGCATGGCGCGCCAGCGCTTATGACGAAMACATGGCGCAAAAGTATGCTC	6556
O	527	acgtttccaaagcaagatgtgttacaagcatccatttttagtgatcgcgaagtaaaaa	586
D	6555	ACCTTGACACAGGTAATCTGGAATATAACACCCCTTACCTGGAACACGCGACAGTTAAAC	6496
O	587	aagtggcagatgattgtgtgttccacaagacgtagtcttgtgtgtgacttaagtgtaac	646
D	6495	GGGTGGGCTACGTGTGTGTGTGCACGACCGAGCGTGTGTGTGCGGTGTGAACATTAAAC	6436
O	647	tgtttaaactgtatataatgaaaaagaaatgynaagaagaaagatgttctgcgttcaat	706
D	6435	TGTTCAAAAACCTGCTGCGCGAATATGAACACCGACGACGAACGCGTCAATGCGAC	6376
O	707	tgaagttaatcgtgttccaaatcatacctaacttttccaacttttggaaatataatttaa	766
D	6375	TGCGAATGATCGGCTCGAAAAGCGGTGCGTTCCTTCAACCTCGTGGCGCATGTGTTG	6316
O	767	cccaagaatcaggtatggtgtgaatccctcctgttgagagaagtaattggttcagtaaat	826
D	6315	CCCAAGTACCGCGATGGGGGATAAACCTTCCCTGTGCGAATGATCGGTCCGTTAAAG	6256
O	827	ctatgattagtgcttaataaaaaaggggaagtagatgttgttattagtttataacaat	886
D	6255	TGATGTTGACAGGCGCTACGAGAAAGGCGCGTGTGAGACAGTTTACATGTCAGCAACAAT	6196
O	887	ttattacaacgatgtcgcgaagaagccagtatgtgaaaaatataatccattaccagaatat	946
D	6195	TTATTAAACACCATGTCTCAAGTTCACACCATCAGCACAGCATGTGCGCTTACCGGCATCAG	6136
O	947	ataatgtgaattaggcgaagaagaacaagtttggatataattagaacactatgta	1006
D	6135	ATGATGATGATCT-----GAAACATAAATTCCTGGGATTACTGTGACAAACCCATCCGA	6082
O	1007	aagattattagaataattatgttgcgttattagtaaatcgaagtttatcaagcaagcaag	1066
D	6081	AGGCGTGTGCGATACCCCTGCGTCCGTTATGTGGAATTCACGTTTATCAGGCGCTGCG	6022
O	1067	ttgaaaaacttgcttcgacgaagccgctcgaaatggtcgcgatgaagaacgaacagata	1126
D	6021	TTGTAAAACTGGCGACGAGACAGGCGCGCCGTCGATGTGTCATGAACCCCGACACGACA	5962
O	1127	acgcaaggtaacttaattatgtagttagttagttatatataaagaagctcgttaagcaagta	1186
D	5961	ATGGCGGACCTGATTAAGAAGCTGCACTGTGTATACAAAGACTGTCATGAGGCCACGA	5902

QY	1187	ttcaaatgaatgaatgaatgtaattgttgcgcggtgcgcgcgaatttaacaa-----a	1237
Db	5901	ttactcTAGGAAACTCACCgGATGCTCGGGGGCCGCCGGTTAAACAGTTATTTCCG	5842
QY	1238	tagagatcggtaatgagcaactggaanaattgtacaaatcactcggltgcgttatgact	1297
Db	5841	TGAGAGATTAAAGATGGCTACTGGAAAGATTGTCCAGTAATGGCGCCGCTAGTTGAGCT	5782
QY	1298	tgaattccacaagaatgacgataccaaaagtatatatgacgtt-----aatgttga	1348
Db	5781	CGAATTCCCTCAGGATGCGCTACCCGCCGCTGTACATGCTCTTGAGTGTCAAAAAGTTAA	5722
QY	1349	aacaggttagaacttgaattgaattcaacaacaatlaagtggtgtgtgtgtgtgtatgc	1408
Db	5721	TGAGCGCTGTGGCTGGAAAGTTACAGACGCTCGCGCGGCTATCTGTACGTACCATGCG	5662
QY	1409	aatggatcatctgtatggaataaaacgcggtttaaagcgtlaacaaatacgaataacccat	1468
Db	5661	AATGGGTTCTCTCGAAGGCTCTGCTCGCGGCTGTGGATGTAAAGACCTCGAACCCGAT	5602
QY	1469	tctcttccagtgggaaacgaataattggttcgtatcatatgaactggtgtgaacaaat	1528
Db	5601	TGAAGTCCCGTAGGTAAAGCAGCACTCTGGCGCTGTATCATGAACTGACACTGGGAACCGGT	5542
QY	1529	cgaatgaacgaattgaatcgtgcgtgcagaagaagatgtgctatctaacgcgtgcgcacaag	1588
Db	5541	CGACATGAAAGGCGGAGATCGGTGAAGAAAGCGTTGGCGATTCACCGGCGACACTTTC	5482
QY	1589	ttatgaagaacacatlaacacagtaactctttagaaacgggaattaaagtatcgaact	1648
Db	5481	CTACGAAGAGCTGTCAAACTCTCAGGAACCTGCGTGAACCCGGTATCAAAAGTTATGACCT	5422
QY	1649	agtttgcggttgcgaaaagggtgaaagttaagtattatcgtgtgtgcggtgtgcgttaa	1708
Db	5421	GATGTGTCGGTCTGATGAAGGGCGGTAAAGTTGGTGTGTTGGTGGTGGCGGTGTAGGTAA	5362
QY	1709	aaccgtcaaatgaatgaatlaatccgttaacatcgcaattgagcactgaagttaactcgt	1768
Db	5361	AACGTTAAACATGATGAGAGCTCATTTCTGTATACATCGCATCGAGCACTCCGGTACTCTGT	5302
QY	1769	ctttgcgggggtagtgtagcgtgcgtgaaagtaaacgaactctatcatatgatagaaga	1828
Db	5301	GTTTTCGGGGCGTAGGTGAACGTACTCGTAGAGGTATACCACTTCTCCAGAAATCCGGA	5242
QY	1829	ctctaacgtatagaataagtgctcctgttatatgttaaatgaacaggagcaaccggttaa	1888
Db	5241	CTCCAAAGTTATCGAACAAAGTATCCCTGGTGTATGGCCAGATGAAGACGCGCCGGGAAA	5182
QY	1889	ccgtttacggttggatlaacagcgttlaactaactgcgggaaanaattcgttgaagtlcg	1948
Db	5181	CGGTCTGGCGGTGTGCTGACCGCGTCTGCATATGCTGTGAAGAAATTCGCTGACGAAGGTGCG	5122
QY	1949	tgatgtcttatcttcgttgtataa 1972	
Db	5121	TGACGTTCTGCTGTCTGTGACAA 5098	

RESULT	6
LOCUS	ECUNC
DEFINITION	E. coli origin of replication oric and genes gid, unc, Ecourf-1 and gms.
ACCESSION	X01631
VERSION	X01631.1 GI:43256
KEYWORDS	Inverted repeat; origin of replication; transposable element.
SOURCE	Escherichia coli.
ORGANISM	Escherichia coli
REFERENCE	Bacteria: Proteobacteria; gamma subdivision; Enterobacteriaceae; Escherichia.
AUTHORS	1 (bases 1 to 14526)
TITLE	Walker, J. E., Gay, N. J., Saraste, M. and Eberle, A. N.
	DNA sequence around the Escherichia coli unc operon. Completion of

the sequence of a 17 kilobase segment containing asnA, oriC, unc, glms and phos
Biochem. J. 224 (3), 799-815 (1984)

JOURNAL
MEDLINE
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
SOURCE

2 (bases 1 to 14526)
Messer, W., Hartmann-Kuehlein, H., Langer, U., Mahlow, E., Roth, A.,
Schaper, S., Urmonet, B. and Woelker, B.
The complex for replication initiation of escherichia coli
Chromosoma 102, 1-6 (1992)
Location/Qualifiers
1. 14526
/organism="Escherichia coli"
/db_xref="taxon:562"
1. 28
/note="16k protein"
/codon_start=2
/transl_table=11
/protein_id="CAA25772.1"
/db_xref="GI:809693"
/db_xref="SWISS-PROT:P03817"
/translation="GSVNLK"
65. 296
/note="min. region of oric"
343. 348
/note="pot. promoter -35 region"
367. 372
/note="pot. promoter -10 region"
395. 398
/note="pot. Shine Dalgarno region"
407. 2293
/note="gida protein"
/codon_start=1
/transl_table=11
/protein_id="CAA25773.1"
/db_xref="GI:43257"
/db_xref="SWISS-PROT:P17112"
/translation="MYDPEDVYIIIGGSHAGTEAAMAARMGQTLITLTHNIDPLG
MSCNPAIGIGKGLVKEVDALGMLAKAIDAGIOFRILNASKGPAAVATRAQADRY
LYROAVRTALENQPMIFQAVEDLIYENDRYGAVTQGLKFRKAVLVLTGFFLD
GHIHGLDNYSGGRAGDPSIPLSRRLRELPLRVGLKGTGPRIDARTIDFVLAQO
HNDNRPVSPMGNSOHPQOVCYTIHNEKTHDYIRSNLDRSPMYAGVIGVPRY
CSIEKVRFRADRNHOJFLEPBGTSNEIYPNGISTSLPDYMOQIYRMOGMEA
KIVRQYALEYDFDRDLPTLESKFIFGLFPAGQINNTGTGEAAAGLLAGLAA
RISADKEGMAPASQAYIGLVLDLCTGCTKEPRFTSRAEYRLMLREDNARLITE
IGRELGLVDERNARNEKLENIERERQRLKSTWVPSAAAEVNAHLTAPLSRAS
GSDLRPEMTYERKLTLTPEFAPALTEQAAEQVEIQYKEGYIARQODIEKQIREN
TILPRLDXYRQYSGLSNEYIAKNDHKPASIGASRISGVTPAASITLLVWLKQGL
NRSA"

repeat_unit
RBS
CDS

2305. 2330
/note="pot. hairpin structure"
2346. 2349
/note="pot. Shine Dalgarno region"
2357. 2980
/note="gida protein"
/codon_start=1
/transl_table=11
/protein_id="CAA25774.1"
/db_xref="GI:581247"
/db_xref="SWISS-PROT:P17113"
/translation="MNLKSLIKDAGISTLTDHOKNOLIAVYMLKRMKNKAYNLTSVR
DPNEMLVRLHLDSTIVAPYIQRERFDVGTGPGRLPGIPLSTIVRPEAHFTLDSIGRV
RFLROYQHELKLEIEPVSVEEPSEPPDVISRAFASLNDVSWCHHLGBOGR
FYALKGMPEDDELALPEEYQVESVVKIQLPALDGERHLVVIKANKI"

promoter
misc_feature
CDS

3474. 3479
/note="pot. -35 region"
3500. 3505
/note="pot. -10 region"
3512
/note="pot. transcription start site (unc)"
3585. 3977
/note="uncl protein"
/codon_start=1
/transl_table=11

/protein_id="CAA25775.1"
/db_xref="GI:581248"
/db_xref="SWISS-PROT:P03808"
/translation="MKNVMSYSLVSRVARKLLVOLLVYIASGLFSLDPEWGSVA
ISGLAVFLPVLVLFMIFAMRHQAATPAKRAMTFARFAGEAKVLAAMLVLLVALAVLK
AVFLPLIVTWVLVVOILAPAVINKG"
3979. 3981
/note="pot. Shine Dalgarno region"
3986. 4801
/note="uncb protein"
/codon_start=1
/transl_table=11
/protein_id="CAA25776.1"
/db_xref="GI:43260"
/db_xref="SWISS-PROT:P00855"
/translation="WASENMTPODYIGHLLNNLOLDRFSLDVPONPAPFTINID
SMFSSVLGLFLVLFPSVAKKATSGVGRKOTFELIYVIGNSKDMYGSKSLIA
PLALIFWVFLMNLMDLLPIDLPIYIEHVLGDPALAVYSADYNATLSALGFEIL
ILEFSIKMKGIGGFKELTLPFNHMAIPLNLILEGVSLLKSPVSLGRLEFGMYAG
ELIFLIGLPLPMSOWILNVPMAIFHILITLQAFIFMVLIYLSMASEB"
4838. 4841
/note="pot. Shine Dalgarno region"
4848. 5087
/note="uncb protein"
/codon_start=1
/transl_table=11
/protein_id="CAA25777.1"
/db_xref="GI:43261"
/db_xref="SWISS-PROT:P00844"
/translation="MENLMDLILMAALVPMGLAIGAIGIGIGKFLBGAARQPD
LPLRLTQFFIVMGLVDAILPMIATVGLGYMFAVA"
5146. 5149
/note="pot. Shine Dalgarno region"
5149. 5619
/note="uncf protein"
/codon_start=1
/transl_table=11
/protein_id="CAA25778.1"
/db_xref="GI:581249"
/db_xref="SWISS-PROT:P00859"
/translation="MNLNTIIGQAIAFLVLFQMKYVPPMLAAIEKREKIADGL
ASAEARHKLDLAKASATDQLKAKAEAVIITBOANKRSQULSDAKAABDEBKIV
AOQAQIEIAERKARELRKQVAILVAAGETIERSVDEANSLDVKLAEV"
5619. 5622
/note="pot. Shine Dalgarno region"
5625. 5628
/note="pot. Shine Dalgarno region"
5634. 6167
/note="uncl protein"
/codon_start=1
/transl_table=11
/protein_id="CAA25779.1"
/db_xref="GI:43263"
/db_xref="SWISS-PROT:P00831"
/translation="MSEFIVARPYAKAFDFEAVERHOSVERMODLAFPAEYTKNKM
AELLSCALAPETLAESEFIAVCGEOLDENGONLIRVAENGRIALADVLEQFIHLAV
SEATAVDVISAALSEQOLAKISAMERKLSRKVLNCKIDKSVNAGVIRAGDMVI
DGSVRGTERLADVLOS"
6175. 6178
/note="pot. Shine Dalgarno region"
6180. 7721
/note="uncl protein"
/codon_start=1
/transl_table=11
/protein_id="CAA25780.1"
/db_xref="GI:43264"
/db_xref="SWISS-PROT:P00822"
/translation="MQLNSTEISLILKQRIAGPNVNSEAHNECTIVSDDGIYRIHGL
ADCMOGEMISLPGNRVAILNLERSDVGAVMGPYADLAGMKVYKGTGRIIEVPVGG
ILGRVYVIGAPIDGGLDHPDSAEALAVIEROSVDOPVOTGSKAVDSMIPIG
RGQRELITGRQGTKRLAIDAIINODSICIKIYVAGKASRTISNVKLEBHGAL
ANTIVVATASESALQYIARMPVALLGERTFRDGEALILTYDLSQAAYROIISLL
LRPPGREFAPGVDFVILHSRLERAAVNAVEYEAFTKGKGTGTSLTALPIETQIA

GDVSAFVPTNVIITDQIFLETNLNFNAGIRPAPVPGISVRGGAOTKIMKLKSG
IRRLAQYRELAASFOSAPSDLDADATRKOLDKQVTELLKOKXAPMSVAQOSIYLF
AERGLADVELSKISGFPAALAYVDHBMIMQEIINGTGSYNDIEGKILKGIIDSK
ATOSM"

repeat_unit
7726..7753
/note="pot. hairpin structure"
7761..7764
/note="pot. Shine Dalgarno region"
7772..8635
/note="uncd protein"
/codon_start=1
/transl_table=1
/protein_id="CAA25781.1"
/db_xref="GI:43265"
/db_xref="SWISS-PROT:P00837"
/translation="MAGAKETRSKIASVQNTQKTKAMENYVASKMRKSODRMAASRP
YAEIMRVYIGHLAHLNLEKHPYLEDKRVKGLVYSTDGLCGNINIFKLLE
MKWTMDGVOCDLAMIGSKVSPNSGVNVAQVMDGAVTGMSELIGPVKMLQAY
DEGRDLKLYISNFIYNTMSQVPTISOLPLPADDDDKHRSMDYLYEPDKALDPT
LIRREVSQYQGVENVLASBQAARVMKMATNGSLIKELQLYVNKAQASITQE
LLETVSGAAV"

RBS
8652..8655
/note="pot. Shine Dalgarno region"
8662..10044
/note="uncd protein"
/codon_start=1
/transl_table=1
/protein_id="CAA25782.1"
/db_xref="GI:899257"
/db_xref="SWISS-PROT:P00824"
/translation="MATGIQVYIGAVYDVEFPDQAVPRVYDALEVOGNRILEYEO
QULGGIVRTIAMSSDGLRGLDKLEHIEVPKATIGRIMNYLIEPVDKKEI
GEERMAIHRAPSEYELNSQELLEIKYIDLMCPRAKRGKGLGSAVGCTVNM
MELRNATIHSGISVFAGVGERREGDHEHTDSNVLDKSLVVGQMPENRL
RVALTGLMAERKEDRDLVLEVDNIYRTLAETVSALLRMPSAVGQPTLAEM
GVLOERTSTRTGSIYSQAVYVADLTSPATTAHLDAFVLSQIOLSLIYPA
VDPIDSTRDLPLVGOEHDTARVOSIILQRYOELKDIILAILGMDLSEEDLKYA
RARKIORFLSOPFEVFAEFTGSPGKYVSIKMTINGFIMGEDHLEQAFYVVGSI
EEAVERAKKT"

RBS
10057..10059
/note="pot. Shine Dalgarno region"

Query Match 41.2%; Score 812; DB 2; Length 14526;
Best Local Similarity 64.7%; Pred. No. 5.3e-149;
Matches 1296; Conservative 0; Mismatches 660; Indels 48; Gaps 4;

11 ttggctcagtcgctcgagatcttctatgacgacgtaagcaattatcacatgtaaa 70
|||||
7395 TTCTCTCAGTTTGATCGATCGACCTTGACATGCAACAGTAAGCAGCTTGACCA 7454
|||||
71 aaggaactgaattatgaacaaacaaataactctcgttatctgtgacacaaagca 130
|||||
7455 AAAGTACCGAACTGCTAAACAGAAACAGATGCGCGATGTCGTTGCGCAGCAGTCT 7514
|||||
131 ttatgttatttgcagtagcttgccttacttaagaagaagtgacttgcatttgt 190
|||||
7515 CTGGTTCGTTGCGAGCAAGCGTTCACCGGCGGAGTGTGAACCTGCAAAATTTGGC 7574
|||||
191 tcaattgaatcaagcaatttgcagtagtaagtaacataactgctgattttagtcg 250
|||||
7575 AGCTTCGAAGCGCTCTCTGCTTACGTCGACGCTGATCAGCGCTCTGTCGAAAG 7634
|||||
251 ttaaccgaatctggaatataatgaatgaatgaagtagtcaatgaaggaatttgat 310
|||||
7635 ATCAACCGAACCGGTGGCTTACACGACGAATTCGAAGCGCAAGCTTAAGGCAATCCTCAT 7694
|||||
311 agcttcaagaagaacagtgctgtta-----agttaacct 346
|||||
7695 TCGCTCAAGACCAACCAATCTCGTAAGCTGCGGCGCTTGCTTAAGGCAAGCGCGCAAG 7754
|||||
347 ttaaatgtagagacaatagcgagtgtaagaagataagaacaacaaatcgcgagtgtaa 406
|||||
7755 GCATTGAGAGAGAGCTCTATGCGCGCGCAAAAGATATAGTAAAGATCGCAAGCGTCC 7814
|||||

407 aagtagacacaaaataacttaagaagcagtggaatggttgcctcgcgaataatgcgtaa 466
|||||
7815 AGAACACCGCAAAAGATCTACTTAAGCGATGGAGATGCTGCGCCCTTCCAAATGCTTAAT 7874
|||||
467 cgaagaacgcagatgctcttcttcacgccccttattcagaacaatacgaatcagtagtaagc 526
|||||
7875 CGCAGATCGCATGCGCGGCGCAGCCGTCCTTATGACAGAAACCATGCGCAAGATGTGGTC 7934
|||||
527 acglttccaagaacagatggttgaacagatcattttagtgaatcgcgagtaaaaa 586
|||||
7935 ACCTTGCACACCGTATCTGCAATATTAAGCACCTTACTGTGAACACCGCGCAGCTTAAC 7994
|||||
587 aagtggaatgatttgcctcagacagatcggtgcttgcgttgcgttaaacgtgaact 646
|||||
7995 GCGTGCGCTACCTGGTGGTGTGACCGACCGCGTTTGTGCGGTGTTGAACATTAAC 8054
|||||
647 tgtttaaactgtagtaataatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaat 706
|||||
8055 TGTTCAAAANAACGTCTGCGGGAATGAAGACCTGAGACCGACAAAGCGCTCAATGCGAAC 8114
|||||
707 tgaatttaactggttctaataatcattcaacttttccacttcttgggaattaaatttaa 766
|||||
8115 TCGCATATGATCGCTCGAAAGCGCTGCTCTTCATCTCCGTGCGCGCAATGTTGTG 8174
|||||
767 cccaagattcagtagtattgtagtactccctcgttgcagtagtaattggtcagtagc 826
|||||
8175 CCCAGGTACCGCGCATGGGGGATTAACCTTCCCTCCGAACTGATCGGTCCGTAAG 8234
|||||
827 ctatgattgattctataaaaaaagggaagtagatggttgcgttatttgaatttaacaat 886
|||||
8235 TGATTTGACGCGCCACGACGCAAGGCGCTGTCGACAGCTTACATTTGACACCAAT 8294
|||||
887 ttattaacagatgctgcgaanaaagcagtagtggaaataatccattaccagaattag 946
|||||
8295 TTATTAACACCATGCTCAGGTTCCGACCATGACGACCTGCTGCCGTTACCGCATAG 8354
|||||
947 ataagtagaattagcgaagaagaagaagaagttgagattatataatgaacactgtagcga 1006
|||||
8355 ATGATGATGATCT-----GAAACATTAATCCTGGATTAACCTGTACGAAACCGATCCGA 8408
|||||
1007 aagtagtattgaataattattggttgccttattgaatctcaggttatacagaagcag 1066
|||||
8409 AGCGTTGCTGATATACCTGCTGCGCTGTTATGTGATGATAGTTTAAACAGGTTATTGCG 8468
|||||
1067 ttgaacacttgccttgcgaagcagcgtcgaaatggtgcgaatgaagaagcagaagata 1126
|||||
8469 TTGAATAACCTGCGCAGCAGCAGCGCGCCGCTATGTGCGATGAAGCCGCGACCGACGA 8528
|||||
1127 acgcagtagaacttaattatgtagtgaagtagtctataaagaactgcgtcaagaagta 1186
|||||
8529 ATGGCGGCGCTGATTAAGAGCTGCAATTTGTATACAAACAACTGCTGAGGCGACGA 8588
|||||
1187 ttacaataatgaatgaatgaatggttgcggtgcgagcagaatgaacaa-----a 1237
|||||
8589 TTACTCAGGAACCTACCGAGATCGTCTCGGGGCGCGCGGTTTAAACAGGTTATTGCG 8648
|||||
1238 tagagatcgtgtaattgcgaactggaanaaatgtacaaatcattcgttgcgttattgcgt 1297
|||||
8649 TAGAGAGATTTAAGAGGCTACGAGGAAGATTGTCAGGTAATCGCGCGCGGATGTGAGAGT 8708
|||||
1298 tgaattccacaagaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaat 1348
|||||
8709 GCAATTTCCCTCAGATGCTGCTGACCGCGCGCTGACAGTCTTGGAGTGCAAAATGTGTA 8768
|||||
1349 aacaggttagtacttgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaat 1408
|||||
8769 TGAGCGTCTGATGCTGGAAGTTACACACAGCTGCGGGCGGTATCGTAGACATACATGCG 8828
|||||
1409 aatggatcattcgttagtgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaat 1468
|||||
8829 AATGGGTTCTCCGAGGCTGCTGCGGCTGATGATTAAGATGATTAAGATGATGATGATGATGAT 8888
|||||

QY	1469	tctctgtccagttggaacgcgaacaacatgggtcgtatcatatgaagatattgggtgaacaat	1528
Db	8889	TGAAGTTCCTCCGGTAGGTAAACGCATCTCGGGCCGATATCATTAACGTAATGGGTGAACCGGT	8948
QY	1529	cgaatgacaaagtgaaatcgtgtgcagagaagaattggtctcatcaccgttcgcaccacaag	1588
Db	8949	CGACATTAAGGCGACAGATCCGTTGTAAGAGACGTTGGGCGCATTCACCGCGACAGACCTTC	9008
QY	1589	ttatgaagaacaatctaacagtaactgaactttagaacaacgggaattaaagtatcgactt	1648
Db	9009	CTACGAGAGAGCGTCAACACTCTCAGGAACATGCTGGAACACCGGTATCAAGTTATCGACCT	9068
QY	1649	agtttgcgtgttcgcgaagaagggttaagtaagttatttcgtgtgttcggtgtgcgttaa	1708
Db	9069	GATGTTCCTCCCTTCCCTTAAGGCGCGTTAAAGTTGTCGTGTGTCGGTGGGTGAGGTAA	9128
QY	1709	aaccgtaaatatgtatgtgaatcaatccgtaacatcgcaattgaacactcgaagttactctgt	1768
Db	9129	AACCGTAAACATGATGAGNCTCATTTCCGTAACATGCGCATCGACACTCCGGTTACTCTGT	9188
QY	1769	cttcgcggggatgaagttgagcgtgcgtgtgaaggttaacgactctatcatgatgaaga	1828
Db	9189	GTTTGGCGGGCTAGGTGTAACGTAATCGNAGGGTAACGACTTCAACACGAATGACCGA	9248
QY	1829	ctctaacgattatagataaagtgtctctctgtttatgtgtcataatgaacagccacaaggtaa	1888
Db	9249	CTCCAACTGTATCGCAACAACTATCCCTGGTGTATGCGCATGTAACGAGCCGCGGGA	9308
QY	1889	ccgtttaagttgtgcattaaagagcttaacatcgcggaataatccgttgtaagttcg	1948
Db	9309	CCGTCGTGGCGGTGCTCTGACGCGGTCTCAACCATGCTGATGAAATTCCTGACGAAGTCG	9368
QY	1949	tgaatcttatctctctgttgaata	1972
Db	9369	TGACGTTCTGCTGTCTGTTGACAA	9392
RESULT	7		
ECOWM82/c			
LOCUS	ECOWM82	136254 bp	DNA
DEFINITION	E. coli; the region from 81.5 to 84.5 minutes.		
ACCESSION	L10328		
VERSION	L10328.1	GI:290484	
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS	1 (bases 1 to 136254)		
TITLE	Burand,V., Plunkett,G. III., Daniels,D.L. and Blattner,F.R.		
JOURNAL	DNA sequence and analysis of 136 kilobases of the Escherichia coli		
MEDLINE	genome: organizational symmetry around the origin of replication		
COMMENT	genomics 16 (3), 551-561 (1993)		
	93315143		
	This sequence was determined as part of the E. coli Genome Project		
	(Frederick R. Blattner, director) at the University of		
	Wisconsin-Madison. Supported by award HG00301 from the NIH Human		
	Genome Project. A preliminary report was presented at the Hilton		
	Head meeting Genome Sequencing and Analysis IV, September 26 -		
	30, 1992. The entire sequence was independently determined from E		
	coli MG1655; overlaps and conflicts with other sequence		
	determinations are annotated. The end of this entry overlaps the		
	start of the entry ECOWM85 (M87049) by the six bases of an EcORI		
	site.		
	Data kindly submitted in computer readable form by: Guy Plunkett		
	III		
	Laboratory of Genetics		
	University of Wisconsin		
	445 Henry Mall		
	Madison WI 53706		
	USA		

Phone: 608-262-2534
Email: ecoligenetics.wisc.edu
Fax: 608-263-7459
[1] Authors request hold until publication
location/Qualifiers
1. 136254
source
Escherichia coli

```

misc_feature      complement(<1. .669) "
                  /note="X06036; ECOFG(1. .>669)"
misc_feature      1. .11607
                  /note="EC14-52 (lambda clone)"
CDS               complement(<1. .544)
                  /gene="fpg"
                  /note="alternate gene name mutH"
                  /codon_start=1
                  /transl_table=11
                  /product="formamidopyrimidine-DNA glycosylase"
                  /protein_id="AAA61988.1"
                  /db_xref="GI:643609"
                  /translation="MPQLPEVETSRRGIEPHLVGATILAAVNVNGKLRLPPESEIYRL
SDQVLSVORAKYLLLELPEGMIIHLGSGSLRIILPELPEKHDHVDVNSGKV
LRVDDPRFGAMLTKELEGHNVYTLHGPEPLSDDENGEYLIHQCAKKTAIKPWLMD
NKLIVGNGNIYASLSFAAGI"
gene              complement(1. .544)
                  /gene="fpg"
misc_feature      complement(526. .1290)
                  /note="J01677; ECOFPMBC(1. .764)"
promoter          complement(557. .586)
                  /note="promoter-like sequence; promoter matrix score of
44"
misc_difference   596. .601
                  /note="AAAAA in X06036; AAAAA in J01677 and here"
misc_difference   611
                  /note="G in J01677; A in X06036 and here"
misc_difference   615
                  /note="G in J01677; A in X06036 and here"
misc_difference   618. .623
                  /note="TTTTTTT in X06036; TTTTTT in J01677 and here"
gene              complement(642. .809)
                  /gene="rpmg (CG Site No. 236)"
CDS               complement(642. .809)
                  /gene="rpmg (CG Site No. 236)"
                  /codon_start=1
                  /transl_table=11
                  /product="50S ribosomal subunit protein L33"
                  /protein_id="AAA61989.1"
                  /db_xref="GI:290486"
                  /translation="MAKIRREKIKLYSAGCHYTTTKNKRTPKELLEKFPDVR
OHVYKEAKIK"
gene              complement(830. .1066)
                  /gene="rpmB (CG Site No. 240)"
CDS               complement(830. .1066)
                  /gene="rpmB (CG Site No. 240)"
                  /codon_start=1
                  /transl_table=11
                  /product="50S ribosomal subunit protein L28"
                  /protein_id="AAA61990.1"
                  /db_xref="GI:290487"
                  /translation="MSRVCOVYGRKRPVGNNSHALNATKRRLPNLHSHRFVWESEKE
RVTYLRVSKMGKRVYIDKKGIDIVLAEELARGKY"
misc_difference   902
                  /note="G in J01677; A here"
misc_difference   1128
                  /note="TC in J01677; C here"
misc_difference   1153. .1155
                  /note="T in J01677; TAT here"
misc_feature      complement(1168. .1819)
                  /note="X63366; ECOFADC(1. .651)"
promoter          complement(1212. .1240)
                  /note="promoter-like sequence; promoter matrix score of
77"
                  complement(1283. .1957)
gene

```


QY 707 tgaagtaatcggcttaataatcatacacttttccaatcttctgggaatlaaatttaa 766
 Db 107314 TCGCAATGATGCGCTGGAAGGCGTGTCTTCAACTCCGGGCGCAATGTGTTG 107255
 QY 767 cccaagaattcagatggtgatactccctcgttgagcaagtaattggttcagcaatt 826
 Db 107254 CCGAGGTACCGGCATGGGGGATACCTTCCCTGCCGAACATGATCGCGGTAAAG 107195
 QY 827 ctatgattgacttataaaaaaggggaagtaatgtgtgtattagttatatacaaat 886
 Db 107194 TGATGTTCAGGCGCTACGAGAGGCGCTGTGAACAAGTTTACATTGTACGACAAAT 107135
 QY 887 ttattacaagatgctgcgaagaagccagatttggaataaatccattccacgaattag 946
 Db 107134 TTTATTACACCATGTCTCAGTTCCGACATGACCGCGTGTCCGCTTACCGGCATCAG 107075
 QY 947 ataataatgaatagcggaagaagaacagtttgggattatattcgaacctgata 1006
 Db 107074 ATGATGATGATCT-----GAAACATTAATCTGGGATTAACCTGTACGAACCCGATCCGA 107021
 QY 1007 aagttattatagaataattatttgctgtattatagaatctcaagttatcaagcag 1066
 Db 107020 AGGCGTTGCTGAGACCTCTCTGCTGTATGTGCAATCTCAGATTATCAAGGCGCTGG 106961
 QY 1067 ttgaacacctgtctctcgaagaacgcgcgaatgctgcacatgaagaagcaagata 1126
 Db 106360 TTGAAAACCTGCGCACAGGACGAGCGCCCGGTATGATGGGATTAAGCCGCGACCA 106901
 QY 1127 acgcaggaacttaataatagttagttagttatgctatacaaaagctcgtcaagca 1186
 Db 106900 ATGCGCGACCGCTGATTAAAGAGCTGACAGTGTATACAAAGCTCTGACGCCAGCA 106841
 QY 1187 ttacaatgaatataaagaattgttcggcggtgcagcaacaaat-----a 1237
 Db 106840 TTATCTCAGGAACCTACCGAGATGCTCTCGGGGCGCGCGGTTTAAACAGATTATTTCG 106781
 QY 1238 tagagatcggtaatggaacatggaataatgtaacaatacgcgtggtgattgacgt 1297
 Db 106780 TAGAGGATTTAAAGTGGCTACTGGAAGATTGTCCAGGTAAAGCGCGCGATTTGACGT 106721
 QY 1298 tgaattccacaagaatgcagtaacaaagtataagatgcct-----aaatgttga 1348
 Db 106720 CGAATTCCTCCTAGATGCGGACGCGGCTGTACGATCTCTTGAGTGCATAAATGATA 106661
 QY 1349 aacaggttacttgaagttaacaaacaaataggtgtgtgtgactcgtatgc 1408
 Db 106660 TGAGCGTCTGCTGCTGGAAGTTTACGACACACTCGCGCGGTATCGATACCATGCG 106601
 QY 1409 aatggatcatctgattgaataaagcggttaagcgtaacaatacgaataacccaat 1468
 Db 106600 AATGGGTTCTCCGACGAGTCTGCGTCTGGATGTAAAGACTCGAACCCGAT 106541
 QY 1469 ttctgtccagttgggaacgaacaaacttggctcgtctatgaacgattgggtgaacaaat 1528
 Db 106540 TGAATCTCCCGGTAGGTAAAGCGACTGTGGCGTATGAACTGACTGGGGAACCGGT 106481
 QY 1529 cgaatgaagcaatgaataacggttcagaaagaatggtctatccacggtgcgcacaag 1588
 Db 106480 CGACATGAAGGCGGATCGGTGAAGAGAGCGTTGGCGGATTCACCGGCGACACTTC 106421
 QY 1589 ttatgaagaacatcaacagttactgaacttttagaaacggaatlaaagttacgaact 1648
 Db 106420 CTACGAAGAAGCTGTCAAACTCTCAGGAAGTGTGGAACCGGTATCAAAATTATGACCT 106361
 QY 1649 agtttgcgttgcgaagaaggggttaagtaagttatcgggtgtggtggtcggttaa 1708
 Db 106360 GATGTGCTCGTCCGTTAAGGCGGTAAAGTTGCTGTTCGGTGTGCGGGGTAGGTA 106301
 QY 1709 aacgctcaatagatgaagtaataccgtcaacatgcaattgacgaactcaagttactcgt 1768
 Db 106300 AACCTTAACATGAAGAGGACTATTCTGTAACATCGCATCGAGCATCCGGTTACTCTGT 106241
 QY 1769 ctctcggggggtgagtcgtaacggtgaagtaacgactctatcatgagtgaaga 1828

Db 106240 GTTGGCGCGGTAGGATGACTCTGTGAGGGAACGACTTTACACGAATGACCA 106181
 QY 1829 ctctaacgtatagataaagtgctctgttatttgtaacaaatgaacagaccacagtaaa 1888
 Db 106180 CTCACAGCTTATGACAAAGTATCCCTGCTGTATGGCCAGATGACAGAGCGCGGAAA 106121
 QY 1889 ccgttacgtgtgcatlaaacaggtctaactatggtcggaataatccgtgataagttcg 1948
 Db 106120 CCGTCTGCGGCTGTCTGTGACCGGCTGTGACCATGCTGAGAAATTCGTTACGAGAGTGC 106061
 QY 1949 tgatgtctattcttctgtgtataa 1972
 Db 106060 TGACGTCTGCTGTGCTGTGACAA 106037

RESULT 8
 AE005605/c
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM

AE005605 11999 bp DNA BCT 21-MAR-2001
 Escherichia coli O157:H7 EDL933 genome, contig 3 of 3, section 224
 of 290.
 AE005605 AE005174
 AE005605.1 GI:12518586

Escherichia coli O157:H7 EDL933.
 Escherichia coli O157:H7 EDL933
 Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 Escherichia.

REFERENCE
 AUTHORS
 1 (bases 1 to 11999)
 Perna,N.T., Plunkett,G., III, Burland,V., Mau,B., Glasner,J.D.,
 Rose,D.J., Mayhew,G.F., Evans,P.S., Gregor,J., Kirkpatrick,H.A.,
 Postal,G., Hackett,J., Klink,S., Boutin,A., Shao,Y., Miller,L.,
 Grodebeck,E.J., Davis,N.W., Lim,A., Dimalanta,E., Potamousis,K.,
 Apodaca,J., Anantharaman,T.S., Lin,J., Yen,G., Schwartz,D.C.,
 Welch,R.A. and Blattner,F.R.
 Genome sequence of enterohaemorrhagic Escherichia coli O157:H7
 Nature 409 (6819), 529-533 (2001)

TITLE
 JOURNAL
 MEDLINE
 PUBMED
 REFERENCE
 AUTHORS

2 (bases 1 to 11999)
 Perna,N.T., Plunkett,G., III, Burland,V., Mau,B., Glasner,J.D.,
 Rose,D.J., Mayhew,G.F., Evans,P.S., Gregor,J., Kirkpatrick,H.A.,
 Postal,G., Hackett,J., Klink,S., Boutin,A., Shao,Y., Miller,L.,
 Grodebeck,E.J., Davis,N.W., Lim,A., Dimalanta,E., Potamousis,K.,
 Apodaca,J., Anantharaman,T.S., Lin,J., Yen,G., Schwartz,D.C.,
 Welch,R.A. and Blattner,F.R.
 Direct Submission
 Submitted (22-OCT-2000) Laboratory of Genetics, University of
 Wisconsin, 445 Henry Mall, Madison, WI 53706, USA

FEATURES
 source

Location/Qualifiers
 1..11999
 /organism="Escherichia coli O157:H7 EDL933"
 /strain="EDL933"
 /serotype="O157:H7"
 /db_xref="taxon:155864"
 /note="enterohemorrhagic"
 complement(112..1482)
 /gene="glnu"
 /note="25228"
 complement(112..1482)
 /gene="glnu"
 /note="25228"
 complement(112..1482)
 /gene="glnu"
 /function="enzyme: Central intermediary metabolism:
 Sugar-nucleotide biosyntheses, conversions"
 /note="Residues 1 to 456 of 456 are 100.00 pct identical
 to residues 1 to 456 of 456 from Escherichia coli K-12
 strain MG1655: B3730"
 /codon_start=1
 /transl_table=1
 /product="N-acetyl glucosamine-1-phosphate
 uridylyltransferase"
 /protein_id="AAG58933.1"
 /db_xref="GI:12518587"
 /translation="MLNNASVILLAGKGRMTSDLPKVLHTLAGKAMQVHIDAA

ELGAHVHLYVGHGGDLKQALKDNDNMVLOAEOIGTGHAMQOAPPEADEDLTLM
YGDVPLISEVTGLRLRDAKPOGGTGLITFKLDDPGYGRITREPKVGRGIVPHKATD
EQROIOEINTGILIANGDAMKRMKLKLTNNNAOGYITTDITLALYQSGREIVAVHPQ
RLSEVEGVNNRLQJLSRLERVOSEQAERKLLAGVLRDPARFDLGTLLHGDEVIDT
NVILLEGNTLGHVRKVTGCVIKNSVIDDCEISPYTVEEDNMLAACTIGFPALRP
GAELLEGAHVNEFMKRAKLGKSKAGHLYTGAEGIDNVNIGAGTITCYDANK
FKTIGDBVFVGSOTQLVAPVYVGKATIAAGTITRVNGENALISVHPQKRGKWR
RPVKKK"

gene
complement(1834. .2253)
/gene="atpc"
/note="5229"
complement(1834. .2253)
/gene="atpc"
/function="enzyme; ATP-proton motive force
interconversion"
/note="Residues 1 to 139 of 139 are 99.28 pct identical to
residues 1 to 139 of 139 from *Escherichia coli* K-12 strain
MG1655: B3731"
/codon_start=1
/transl_table=11
/product="membrane-bound ATP synthase, F1 sector,
epsilon-subunit"
/protein_id="AAG58934.1"
/db_xref="GI:12518588"
/translation="MAAMYHLDVSAEQOMSGIVETQVGSSEGLGIYGHAPILT
AIKGMIRIVQHGHEERIIYLSGILEYQJGNTVYLAOTAIRGQDLDERAKEMARKKA
EEHISSHGDVYQAQASALAKALQLNVLTKRAM"
complement(2274. .3656)
/gene="atpd"
/note="5230"
complement(2274. .3656)
/gene="atpd"
/function="enzyme; ATP-proton motive force
interconversion"
/note="Residues 1 to 460 of 460 are 100.00 pct identical
to residues 1 to 460 of 460 from *Escherichia coli* K-12
strain MG1655: B3732"
/codon_start=1
/transl_table=11
/product="membrane-bound ATP synthase, F1 sector,
beta-subunit"
/protein_id="AAG58935.1"
/db_xref="GI:12518589"
/translation="WATGKIYGVGAVVVDEPPDAPRVYDALFVONGNEITVLYQ
QOLGGIVRTIAMGSSDGLRRGLDYKDELEHPEVYGRATITRINNVIGEPPDMGELI
GEERWAIHRAAPSTEELSNSQELLETSIKVIDLCPKPAKGKVLGEGAGVKTNM
MELIRNIAIEHSGSVFAGVGEREGNDPYHEMDSNVIDKSLVYGOMEPGNRL
RVALTGLTMAKFRDEGDLVLFVDNIYRTYIAGTAEVSALLGRMPSAVGOYPTLAEM
GVLOERTSTKSTGITSYQAVYVPADDLTPSPATTFPHADATVYLSROISIGIYPA
VDPIDSTROLDPLVYGOEHYDTARGVOSIIQRYELDITAILGMDLSEBDKLVA
RAKIQRFLSQPFVAEYFTGSPKYVSLKDTIRFGKIMEGEYDHLPEQAFYVNGSI
EEAVEKAKKL"

gene
complement(3683. .4546)
/gene="atpg"
/note="5231"
complement(3683. .4546)
/gene="atpg"
/function="enzyme; ATP-proton motive force
interconversion"
/note="Residues 1 to 287 of 287 are 100.00 pct identical
to residues 1 to 287 of 287 from *Escherichia coli* K-12
strain MG1655: B3733"
/codon_start=1
/transl_table=11
/product="membrane-bound ATP synthase, F1 sector,
gamma-subunit"
/protein_id="AAG58936.1"
/db_xref="GI:12518590"
/translation="WAGAKETRSKIASYQNTOKITKANEMVAASKRKSODRMAASRP
YAEIMRVYIGHLAGNLEKHYHLEDRDYKRVGLVYSDRGLCGLININFEKLLAE
MKTYDVKVOCDLAMIGSGVSPFNSVGVNVAQYTGMDNPSISELIGPVMYQAOY
DEGRIDKLYIYSNKEINTMSQVPTISOLPLPASDDDLKHSMDYLYEPPKALDIT
LLRRVESQVYGVVENLASBOAARMAVMAKKAATJNGSLIKELOLYVYKAKQASTTQE

LTEIVSGAAV"
complement(4597. .6138)
/gene="atpa"
/note="5232"
complement(4597. .6138)
/gene="atpa"
/function="enzyme; ATP-proton motive force
interconversion"
/note="Residues 1 to 513 of 513 are 100.00 pct identical
to residues 1 to 513 of 513 from *Escherichia coli* K-12
strain MG1655: B3734"
/codon_start=1
/transl_table=11
/product="membrane-bound ATP synthase, F1 sector,
alpha-subunit"
/protein_id="AAG58937.1"
/db_xref="GI:12518591"
/translation="MQLNSTEISELIRKIOAFNVNPSAHNNGTIVSDGVIRHGL
ADCMGEMISLPGNRYATALNERSVGAVMGPYADLABGMKYCTGRITLEVPGRG
ILGRVNTLGAIPIDGKPLDHDGSAVVAIAPVIEROSVDPQVOTGKAYDSMPTIG
RGQRELITGDKQGTALAIIDAIIINORDSGIKCIYVAGKASTISNVKRLDEHGL
ANIIYVAVATASESALQYLAIPYAGCAMEYFRDGEDALIYDLSKQAVYROISSL
LRPPGREAPEGDVYELHSLRLERAARVNAEVEAFTGGEVKGKTSIALPIETQA
GDVSAFVPTNYSITDGOIFLEETNFNGIRPAVNVGISVSGCAQTKIMKLTSGG
I RTALAOYRELAAPSOFASDIDDAITRRKOLDGQKTELLKQXAPMSVAAQOSLYLFA
AERGLADYVEIKSGISFPAALALAYVDRDHAPLMDEINOTGGINDEIEGKLGITDSFK
ATQSW"
complement(6151. .6684)
/gene="atph"
/note="5233"
complement(6151. .6684)
/gene="atph"
/function="enzyme; ATP-proton motive force
interconversion"
/note="Residues 1 to 177 of 177 are 100.00 pct identical
to residues 1 to 177 of 177 from *Escherichia coli* K-12
strain MG1655: B3735"
/codon_start=1
/transl_table=11
/product="membrane-bound ATP synthase, F1 sector,
delta-subunit"
/protein_id="AAG58938.1"
/db_xref="GI:12518592"
/translation="MSERTIYARPYAKAFAFVAGHOSVERMODLAFARATKNEOM
HELLSGALAPETLAEFTAVCGEODENGQNLIRWANGRNALPDVLEDFIHRAV
SEATAEVYVISAALISEQOLAKISAMERKRLSRVKKLCKIDKSYMAGVITIRAGDMVI
DGSVGRLERLADYLOS"
complement(6699. .7169)
/gene="atpf"
/note="5234"
complement(6699. .7169)
/gene="atpf"
/function="enzyme; ATP-proton motive force
interconversion"
/note="Residues 1 to 156 of 156 are 100.00 pct identical
to residues 1 to 156 of 156 from *Escherichia coli* K-12
strain MG1655: B3736"
/codon_start=1
/transl_table=11
/product="membrane-bound ATP synthase, F0 sector, subunit
b"
/protein_id="AAG58939.1"
/db_xref="GI:12518593"
/translation="MNLNATILGOAIAFVLEFLFCMKYVPMPLMAIEKROKETADGL
ASAERAHDLDLAKASATDOKKAAEAOVILIEQANKRRSOLDEAKAEAOERTKIY
AOAOETEAERBARREELRKOVAILAVAGAKIITERSYDEAANSIYDKLVAEL."
complement(7231. .7470)
/gene="atpe"
/note="5235"
complement(7231. .7470)
/gene="atpe"
/function="enzyme; ATP-proton motive force

VERSION	AP002566.1	GI:13363930
KEYWORDS	Escherichia coli O157:H7 (strain:O157:H7, sub_strain:RIMD 0509952)	
SOURCE	DNA.	
ORGANISM	Escherichia coli O157:H7	
REFERENCE	Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae; Escherichia.	
AUTHORS	1 (sites)	
TITLE	Makino,K., Yokoyama,K., Kubota,Y., Yutsudo,C.H., Kimura,S., Yamakawa,K., Ishii,K., Hattori,M., Tatsuno,I., Abe,H., Iida,T., Yamamoto,K., Ohnishi,M., Hayashi,T., Yasunaga,T., Honda,T., Sasakawa,C. and Shinagawa,H.	
JOURNAL	Complete nucleotide sequence of the prophage VP2-Sakai carrying the verotoxin 2 genes of the enterohemorrhagic Escherichia coli O157:H7	
MEDLINE	Genes Genet. Syst. 74 (5), 227-239 (1999)	
REFERENCE	20198780	
AUTHORS	2 (sites)	
TITLE	Ohnishi,M., Murata,T., Nakayama,K., Kuhara,S., Hattori,M., Kurokawa,K., Yasunaga,T., Yokoyama,K., Makino,K., Shinagawa,H. and Hayashi,T.	
JOURNAL	Comparative analysis of the whole set of rRNA operons between an enterohemorrhagic Escherichia coli O157:H7 Sakai strain and an	
MEDLINE	Escherichia coli K-12 strain MG1655	
REFERENCE	Syst. Appl. Microbiol. 23 (3), 315-324 (2000)	
AUTHORS	20557356	
TITLE	3 (sites)	
JOURNAL	Yokoyama,K., Makino,K., Kubota,Y., Watanabe,M., Kimura,S., Yutsudo,C.H., Kurokawa,K., Ishii,K., Hattori,M., Abe,H., Iida,T., Yamamoto,K., Hayashi,T., Yasunaga,T., Honda,T., Sasakawa,C. and Shinagawa,H.	
MEDLINE	Complete nucleotide sequence of the prophage VP1-Sakai carrying the	
REFERENCE	Shiga toxin 1 genes of the enterohemorrhagic Escherichia coli O157:H7 strain derived from the Sakai outbreak	
AUTHORS	Gene 258 (1-2), 127-139 (2000)	
TITLE	20564182	
JOURNAL	4 (sites)	
MEDLINE	Hayashi,T., Makino,K., Ohnishi,M., Kurokawa,K., Ishii,K., Yokoyama,K., Han,C.-G., Ohtsubo,E., Nakayama,K., Murata,T., Tanaka,M., Tobe,T., Iida,T., Takami,H., Honda,T., Sasakawa,C., Ogasawara,N., Yasunaga,T., Kuhara,S., Shiba,T., Hattori,M. and Shinagawa,H.	
REFERENCE	Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genomic comparison with a laboratory strain K-12	
AUTHORS	DNA Res. 8 (1), 11-22 (2001)	
TITLE	21156231	
JOURNAL	5 (bases 1 to 280900)	
MEDLINE	Submitted (26-JUN-2000) Ken Kurokawa, Osaka University, Genome	
REFERENCE	Information Research Center; 3-1, Yamadaoka, Suita, Osaka 565-0871,	
AUTHORS	Japan (E-mail:ken@gen-info.osaka-u.ac.jp),	
TITLE	URL:http://www.gen-info.osaka-u.ac.jp/, Tel:81-6-6879-8365,	
JOURNAL	Fax:81-6-6879-2047)	
MEDLINE	genome project.	
REFERENCE	20564182	
AUTHORS	Location/Qualifiers	
TITLE	1. 280900	
JOURNAL	/organism="Escherichia coli O157:H7"	
MEDLINE	/strain="O157:H7"	
REFERENCE	/sub_strain="RIMD 0509952"	
AUTHORS	/db_xref="taxon:83334"	
TITLE	complement(143..616)	
JOURNAL	/gene="ECS4456"	
MEDLINE	complement(143..616)	
REFERENCE	/gene="ECS4456"	
AUTHORS	/note="similar to YIAI_ECOLI YIAI_ECOLI g1 2367245 percent	
TITLE	identity 98 in 157 aa (conserved in E.coli K-12)"	
JOURNAL	/codon_start=1	
MEDLINE	/transl_table=11	
REFERENCE	/evidence-not_experimental	
AUTHORS	/product="hypothetical protein"	
TITLE	/protein_id="BAB37882.1"	
JOURNAL	/db_xref="GI:13363934"	
MEDLINE	/translation="MNISEVDLHKLTVDSPFLGVOYQOLVRDVIPYOMDALNDRIPEA	
REFERENCE	EPSSHAIEFRFAAGLOESEFGYGMVDQDAKWLAEVAMSCQKPDALERTADVITE	
AUTHORS	LIASQCEDGTLNTFTYKAPBERNSNLAECHLTCAGHLTAGVAFQAFQAGVRRGGA	
TITLE	VVCRLADIDSVFGPDESKLHGYPGHPLEIALMLKLYETEPRLATLTNFGVORGA	
JOURNAL	OPHYDDOYERKRGQTSMTYGPAMVMDKAYSQNLALAOQOTAIIGHAVFYVLTMG	
MEDLINE	VAHLARLSHDSKRODCRLTMMNMAQORLYTTGGISQSGSEAFSDVLPNDVYAE	
REFERENCE	SCASGILMFARBMLEMGDSOYADVMEALYNTVYLGSMADGKHFEVNLDPYAEK	
AUTHORS	LKPNIYDHYAKPIROHWRGCAACCPNTARVYTSIGHYLYPREDALYINITYAGSMVEY	
TITLE	PVENSTLRLVSGNTPMOEOYTVAVESQPPVRRHTLALPLPMTCTOPQITLNGEYBD	
JOURNAL	IRKGLHTIREMGEDTLNLTLPMEVRRVRYNPLVRHAGVVALQRCGLVYCLEADN	
MEDLINE	GESLHNLMPLPADPPTPEGKGLFSKLLIDAPGRYQSNBDOPLMHWDSABAKRO	
REFERENCE	TOTLTFIPWFSNANGEGEMRIWNEEKHP"	
AUTHORS	complement(5412..6152)	
TITLE	/db_xref="GI:13363931"	
JOURNAL	/translation="MNRRTADATKTCICRCEVACANSHOENDCALSPDEFISRI	
MEDLINE	RVTKDHSMTTAVACHQCEDAPCANVCISREGHIFVQTRICIGKSCMLCPFG	
REFERENCE	AMEVVSRRKARAIKCDLCWRETPACVVEACPTKALOCMDYEVQRRLRQOPV"	
AUTHORS	complement(860..1675)	
TITLE	/gene="ECS4457"	
JOURNAL	complement(860..1675)	
MEDLINE	/gene="ECS4457"	
REFERENCE	/note="probable transcriptional regulator, weakly similar	
AUTHORS	to hypothetical transcriptional regulator Y1SR [Bacillus	
TITLE	subtilis] g1 31233061sp P40331 percent identity 24 in 276	
JOURNAL	aa"	
MEDLINE	/codon_start=1	
REFERENCE	/transl_table=11	
AUTHORS	/evidence-not_experimental	
TITLE	/product="putative transcriptional regulator"	
JOURNAL	/protein_id="BAB37880.1"	
MEDLINE	/db_xref="GI:13363932"	
REFERENCE	/translation="MLELSMTLPITVONGGLFISRGVGRHSARLESMTLFEVKGRL	
AUTHORS	TIQERERFLVDAGSLILMPRKQVGEDEPPADIKFTWLFHFKAPNDNRMLTHS	
TITLE	VPQQRKVDVQALVALFRQFUSEDEKHQRTPALEFVLLIQQLTRDRLDQAGAG	
JOURNAL	TTLAMKAOGLINTQYHRLSTALARELYCNADYLGRRYRRFHLTLTEALHROVRE	
MEDLINE	AEKLISEARSLTEVALCGFNDVGYFRQIRFKHTGLTPAAMKRRYSKEHNS"	
REFERENCE	1901..3301	
AUTHORS	/gene="ECS4458"	
TITLE	1901..3301	
JOURNAL	/note="probable permease, similar to hypothetical proteins	
MEDLINE	(Salmonella typhimurium) g1 7442781 p I C65167 percent	
REFERENCE	identity 37 in 444 aa, also similar to transport proteins	
AUTHORS	(putative symporters) e.g. Y1CJ [Escherichia coli (K-12)]	
TITLE	g1 2851421 sp P31435 Y1CJ_EC"	
JOURNAL	/codon_start=1	
MEDLINE	/transl_table=11	
REFERENCE	/evidence-not_experimental	
AUTHORS	/product="putative permease"	
TITLE	/protein_id="BAB37881.1"	
JOURNAL	/db_xref="GI:13363933"	
MEDLINE	/translation="MTSPITTTADIARAKKIDDAISVREKIGYGADAGTYITCLIMN	
REFERENCE	FLTFEYTVFGTLTPALVGTFLALNFPALISDPVGVIAIDRTQSMGFRFMQMLAI	
AUTHORS	PIGIIIGLTFVVPDASMGVKIAMAFTGLLSVGTALNVYCARINTMTRHNEVIS	
TITLE	COSMRFVLCGVAGFVSVGLPMVVALFQGNARVYQGVGLCAIVAMFECFPVY	
JOURNAL	REVRPLSTMGFTLREHLAGLRNNDOLIMVMSLLINVRNIRGGVFLTYVIG	
MEDLINE	STGYSLFETVYTPASITIGYSVTSRTPRPTKYVYVNTLLIATLAVMTETPSGA	
REFERENCE	YQTLMLAVILNGVILGFTLLPLHPSLAFADIDYGMKTHVSSGGRNPFNLPFTKLAN	
AUTHORS	ASSAGIISLLETFVAYQGVENQTAASSLGLIAMEITLLPALFHLTLAAGIRCKLNNP	
TITLE	MMSRIATDLRQRHVP"	
JOURNAL	3312..5282	
MEDLINE	/gene="ECS4459"	
REFERENCE	3312..5282	
AUTHORS	/note="unknown, similar to hypothetical protein [Thermotoga	
TITLE	maritima] g1 7452109 p I F72395 percent identity 37 in	
JOURNAL	635 aa"	
MEDLINE	/codon_start=1	
REFERENCE	/transl_table=11	
AUTHORS	/evidence-not_experimental	
TITLE	/product="hypothetical protein"	
JOURNAL	/protein_id="BAB37882.1"	
MEDLINE	/db_xref="GI:13363934"	
REFERENCE	/translation="MNISEVDLHKLTVDSPFLGVOYQOLVRDVIPYOMDALNDRIPEA	
AUTHORS	EPSSHAIEFRFAAGLOESEFGYGMVDQDAKWLAEVAMSCQKPDALERTADVITE	
TITLE	LIASQCEDGTLNTFTYKAPBERNSNLAECHLTCAGHLTAGVAFQAFQAGVRRGGA	
JOURNAL	VVCRLADIDSVFGPDESKLHGYPGHPLEIALMLKLYETEPRLATLTNFGVORGA	
MEDLINE	OPHYDDOYERKRGQTSMTYGPAMVMDKAYSQNLALAOQOTAIIGHAVFYVLTMG	
REFERENCE	VAHLARLSHDSKRODCRLTMMNMAQORLYTTGGISQSGSEAFSDVLPNDVYAE	
AUTHORS	SCASGILMFARBMLEMGDSOYADVMEALYNTVYLGSMADGKHFEVNLDPYAEK	
TITLE	LKPNIYDHYAKPIROHWRGCAACCPNTARVYTSIGHYLYPREDALYINITYAGSMVEY	
JOURNAL	PVENSTLRLVSGNTPMOEOYTVAVESQPPVRRHTLALPLPMTCTOPQITLNGEYBD	
MEDLINE	IRKGLHTIREMGEDTLNLTLPMEVRRVRYNPLVRHAGVVALQRCGLVYCLEADN	
REFERENCE	GESLHNLMPLPADPPTPEGKGLFSKLLIDAPGRYQSNBDOPLMHWDSABAKRO	
AUTHORS	TOTLTFIPWFSNANGEGEMRIWNEEKHP"	
TITLE	complement(5412..6152)	

D	b	223723	CGAATTCCTCAGGATGCCGACCCGCGGTGTACGATGCTTTGGSTGCAGAAATGGTAA	223664
Oy	1349	aacaggttagtacttgaaagtcaacaacaataaggctggtgtgttagttcgcgtgatccg	1408	
D	b	223663	TGAACGTCCTGGTGCTGGAAAGTTCAGCGACGACTCGGGCGGGGTATCTGTGGTCACTCGC	223604
Oy	1409	aatggatcatcattgatgatlaaagaacgcggtttaangctgtaacaaataecgaataccaat	1468	
D	b	223603	AATGGGTCTCTCCGACGGTCTGCGTCCGGCTGTGGATGTAAAAGACCTCGAACACC GGAT	223544
Oy	1469	tctgttcacagtygggaacgaaaacatttgggttcgtatcatatgaacgtatttggtaaccaat	1528	
D	b	223543	CGAAGTCCCCGATAGGTAAAGCGACTCTGGCGCCGATCATGAACGTACTGGGTAAACGGGT	223484
Oy	1529	cgaatgagcaaggtgtaaactcggtgtgaagaagaagatggtctatccaccgtgtcgccacaag	1588	
D	b	223483	CGAATGTAAAGGCGAGAATCGGTGAAAGAGACGGTTGGCGATTCCACCGCACACCTTTC	223424
Oy	1589	ttaatgaagaacactaactaacagtaactgaacctttaagaacggaaattaagaattacgaact	1648	
D	b	223423	CTACGAAGAGAGCTGTCAAACTCTACGGAACACTGCTGGGAAACCGGTAATCAAATTATCGACT	223364
Oy	1649	agtttgtccgttttgcgaagaagggttaagaagttaagtllatccgtgtgtgcgggtgcgtlaa	1708	
D	b	223363	GATGTGTCCTCCCTCCCTAAAGGCGGTTAAAGTGGTCTTGGGTGCTGCGGGTATAGTTAA	223304
Oy	1709	aaccgttaatrtaatgaataaataacacgtlaaacatcogcaatttgaacacactcaagttactctgt	1768	
D	b	223303	AACCGTAAACAATGTATGGAGCTCATCTCTTAACATCGCATGGACATCGACACTCCGGTATCTGT	223244
Oy	1769	ctttgcggaggtaagtgtagcgtlaacgcgtlgaaggtlaaacgactcttatcatatagaatgaaga	1828	
D	b	223243	GTTTGGCGGCGTAGTGTGAAGCTACTCTGTGAAGGGAAGCACTTTCACACGAATGACCGA	223184
Oy	1829	ctctcaacgatatataaagtgtctctctgtttatgtgtccaatgtaaagagccaccaggtaa	1888	
D	b	223183	CTCCAAACGTTATCATATAAGATTCCTCCGTGTATGGCCAAATGAAACAGCGCCGGGAAA	223124
Oy	1889	ccgtttacgtgtgcatcaaacaggcttaactatagtcggaaaaaatccgtatgaagtctg	1948	
D	b	223123	CCGTCGCGGCTGACTGTGACCGGCTCTGACCAATGCTGGAANAATTCCTGTCACGAAGTGC	223064
Oy	1949	tgaatgcttattctcgtgtgataa	1972	
D	b	223063	TGACGTTCTGCTGTTCCTTGACAA	223040

FEATURES	Source
REFERENCE	AF188265.1
LOCUS	AF188265.1
DEFINITION	Salmonella typhimurium strain TA98 ATP synthase operon, complete sequence.
ACCESSION	AF188265
VERSION	AF188265.1
KEYWORDS	GI:6625695
SOURCE	Salmonella typhimurium.
ORGANISM	Salmonella typhimurium Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae; Salmonella.
REFERENCE	1 (bases 1 to 7421) Khm,H.-K., Heo,N.-J., Ghim,S.-Y. and Song,B.-H. Molecular structure of nonastrictonic atp genes encoding ATP synthase in Salmonella typhimurium
JOURNAL	unpublished
REFERENCE	2 (bases 1 to 7421) Khm,H.-K., Heo,N.-J., Ghim,S.-Y. and Song,B.-H. Direct Submission
AUTHORS	Submitted (24-SEP-1999) Dept. of Biology Education, Kyungpook National University, #1370 Sankyuk-Dong, Buk-Ku, Taegu 702-701, Korea
FEATURES	location/Qualifiers
Source	1..7421 /organism="Salmonella typhimurium"

gene /strain="TA98"
/db_xref="taxon:602"
328. .7227
/gene="ATP synthase operon"
gene 328. .720
/gene="atpI"
328. .720
/gene="atpI"
/note="atpI: uses putative GTG translational start site"
/codon_start=1
/transl_table=11
/product="ATP synthase subunit i"
/protein_id="AAFI9355.1"
/db_xref="GI:6625696"
/translation="MKNWSVLSRVNARKLLLVOLLVIASGLFSLKDPKWSA
ISGGLAVLPNLEFMIFAMRHOAHTPAAGRAVMTFAGCAKVLAMLVVALAVLK
AVFLPLVTVVVLVQILPAVINNKG"
729. .1544
/gene="atpB"
729. .1544
/gene="atpB"
/note="atpB"
/codon_start=1
/transl_table=11
/product="ATP synthase subunit a"
/protein_id="AAFI9356.1"
/db_xref="GI:6625697"
/translation="MASENMTPODYIGHLLNNLQDLRTFSVDPQNPDPATFVFINID
SFEFVVGILFELVFRSVARKATGSGKFOPALELVIGVNGSVDMHGSKLLTA
PLATITVWFLMNLMDLLPIDLPIYAEHTLGIPALRVVSADVNTLSALGVFTIL
ILFYSIKKKGIGGFTKELTLOPFNHWATIPVNLIDEGVSLSKFVSIGLRFGMVG
ELFIPLGLPEWMSQWILNVPWAFIHLITITLQAFIMALTIVYLSMASEH"
1591. .1830
/gene="atpE"
1591. .1830
/gene="atpE"
/note="atpE"
/codon_start=1
/transl_table=11
/product="ATP synthase subunit c"
/protein_id="AAFI9357.1"
/db_xref="GI:6625698"
/translation="MENLMDLTYMAAAVMGGLAIGAIGILGKFLGGAARPPD
LIPLRTOFIFYMGLVDAIPMIVAGLGYVMAVA"
1892. .2362
/gene="atpF"
1892. .2362
/gene="atpF"
/note="atpF: uses putative GTG translational start site"
/codon_start=1
/transl_table=11
/product="ATP synthase subunit b"
/protein_id="AAFI9358.1"
/db_xref="GI:6625699"
/translation="MNLNATILGGAIAVLFLFCMKMYPLMAAIEKROKFIANGL
ASAEAHMDLDLAKASATDOLKKAAGVITIEQANKRSQIIDEAFAEEQETKATV
AQAQIEIAERKAREELRKQVAIIAVAGAEKIIERSVDEANSDIVDKIAEL"
2377. .2910
/gene="atpH"
2377. .2910
/gene="atpH"
/note="atpH"
/codon_start=1
/transl_table=11
/product="ATP synthase delta subunit"
/protein_id="AAFI9359.1"
/db_xref="GI:6625700"
/translation="MSEFIIVARPYAKAFPAVEHOSVERMODMLAFAAEVTYKNEOM
AEISGLAIPAEIASFIVACGEOLDENGQNLIVMAENGNLNLPOVLEOFILRAVY
SEATLEUVITIAAALSFOQLAELISAMKRLSRVKNLCKIDKSVMAGVIIIRAGDWI
DGSVGRERLEADVIQS"
2923. .4464
/gene="atpA"
2923. .4464

1067 ttgaaaccttcttctagcaagccgcctgaaatgtgctgcataagcaagcaaga 1126
 1127 acgcagtaacttaataatgagttacagttatgctatacaaaactcgtcaagaaga 1186
 5272 ATGGGCGGAGCGTATTAAAGAGTTCAGATTGATACAAACAACCTGCTGAGGCAAGA 5331
 1187 ttcaaatgaatataatgaatgttgcgcgtgcagcaagcaatttaaca-----a 1237
 5332 TTACTCAGAACTCACCGAGATCGTCTGGGGGCGCGCGCTTTAAACAGGTTATTTCG 5391
 1238 tagagatcgcgttaatgagcaactggaataatgtacaaatcagctggttattgact 1297
 5392 TAGAGATTAAAGATGCGCTACTGGAAGATTGTCCAGGTAATCGGCGCGGTAGTACGT 5451
 1298 tgaatcccaagaatgagtaagcaaaagatataatgact-----aaatgtga 1348
 5452 CGAATTCCTCAGAGATGCGCTACCGCGGCGGTACGATGCTGTGAGTGCACAAATGTGA 5511
 1349 aacaggtttagtacttgaatgacaaacaataatgagttgtgtgtatgctgtatgc 1408
 5512 TGAGCGTGTGGTGGTGAAGTACACACAGCTCGCGCGGTATCGTCCATCCATCCG 5571
 1409 aatggatcattgattgattaaacgcggtttaagcgttaacaaatacgaataaccat 1468
 5572 AATGGATTCTTCGCGGCTTCGCGGCGGTGTGATGTAAGACCTCGAACACCCGAT 5631
 1469 ttctgttcagtgaggaaacaaacatgtgtcgtatcatgaacgtatggtgaaccaat 1528
 5632 CGAATTCCTCAGAGATGCGCTACCGCGGCGGTATCGAAGCGTACGCGGCGGTG 5691
 1529 cgaatgaagaagtgaaatcgtgtgcgaagaagaatgtgtcattacacgtgcgcaacaag 1588
 5692 CGACATGGAAGCGGAGATGCGTGAAGAGAGCGTGGGATTCACCGCGCGACCTTC 5751
 1589 ttataagaacaatcaacagtaactttagaaacgggaataaagtatcgaact 1648
 5752 CTACGAAGAGCTGTAAACTCTCAGGAACCTGCTGAAACCGGTATCAAAAGTTATGACCT 5811
 1649 agtttgcgttttcgaaaggagggaagtaagtaattatcgtgtgtgcgtgtgaa 1708
 5812 GATGTCGTCGCTGCTAAGGGGCGTAAGTTGCTGTTCGCTGCGGCGGTATGAGTAA 5871
 1709 aacgctcaatatgattgaaataatcgcgaatacgcgaattgagcaactaagttactgt 1768
 5872 AACCTTAACATGATGAGGAGCTCATCTGTAACATCCGATCGACATCCGCTTACTCTGT 5931
 1769 ctgtgcgggggtgagtgagcgtgacgcgtggaaggttaagcaacttatcatgagaaga 1828
 5932 GTTTCGGGGCGTGAAGCTACGTCGAGGAGTAACGACTTCTACCAAGAAATGACCGA 5991
 1829 ctctaactatagaataagtgctctgtttatgtaagtaagcaagcaagcaagtaa 1888
 5992 CTCACAGCTTATCGACAAAGTATCCCTGCTGATATGCGCAGATGAACAGCCCGGAAA 6051
 1889 cagttacgtgtgacataacagcgttaactaactgagcgggaaataatcgtgtgaaagtcg 1948
 6052 CGGTCGCGCGTGTGCTGACGCGGTGACCATGCTGAGAAATTCGTTGACGAAGGTGCG 6111
 1949 tgatgtctatcttctcgtgtataa 1972
 6112 TGACGTCTGCTGCTGTTGACAA 6135

RESULT 11
 LOCUS 7881 bp DNA BCT 26-APR-1993
 DEFINITION E.coli atp (unc) operon containing nine genes coding for ATP synthase subunits.
 ACCESSION J01594 J01595 K02181 M25634 V00264 V00313 X00771
 VERSION J01594.1 GI:148131
 KEYWORDS ATP operon; ATP synthetase; ATPase; adenine triphosphate

SOURCE
 ORGANISM
 Escherichia coli
 Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae; Escherichia.
 REFERENCE
 AUTHORS
 TITLE
 Nucleotide sequence of the genes coding for alpha, beta and gamma subunits of the proton-translocating ATPase of Escherichia coli
 Biochem. Biophys. Res. Commun. 103, 604-612 (1981)
 JOURNAL
 MEDLINE
 REFERENCE
 AUTHORS
 TITLE
 Kanazawa, H., Mabuchi, K., Kayano, T., Nouni, T., Sekiya, T. and Futai, M.
 Nucleotide sequence of the genes for F-0 components of the proton-translocating ATPase from Escherichia coli: Prediction of the primary structure of F-0 subunits
 Biochem. Biophys. Res. Commun. 103, 603-620 (1981)
 JOURNAL
 MEDLINE
 REFERENCE
 AUTHORS
 TITLE
 Nielsen, J., Hansen, F.G., Hoppe, J., Friedl, P. and Von Meyenburg, K.
 The nucleotide sequence of the atp genes coding for the F-0 subunits a, b, c and the F-1 subunit delta of the membrane bound ATP synthase of Escherichia coli
 Mol. Gen. Genet. 184, 33-39 (1981)
 JOURNAL
 MEDLINE
 REFERENCE
 AUTHORS
 TITLE
 Gay, N.J. and Walker, J.E.
 The atp operon: Nucleotide sequence of the region encoding the alpha-subunit of Escherichia coli ATP synthase
 Nucleic Acids Res. 9, 2187-2194 (1981)
 JOURNAL
 MEDLINE
 REFERENCE
 AUTHORS
 TITLE
 Gay, N.J. and Walker, J.E.
 The atp operon: Nucleotide sequence of the promoter and the genes for the membrane proteins, and the delta subunit of Escherichia coli ATP-synthase
 Nucleic Acids Res. 9, 3919-3926 (1981)
 JOURNAL
 MEDLINE
 REFERENCE
 AUTHORS
 TITLE
 Kanazawa, H., Kayano, T., Kiyasu, T. and Futai, M.
 Nucleotide sequence of the genes for beta and epsilon subunits of proton-translocating ATPase from Escherichia coli
 Biochem. Biophys. Res. Commun. 105, 1257-1264 (1982)
 JOURNAL
 MEDLINE
 REFERENCE
 AUTHORS
 TITLE
 Kanazawa, H., Mabuchi, K. and Futai, M.
 Nucleotide sequence of the promoter region of the gene cluster for proton-translocating ATPase from Escherichia coli and identification of the active promoter
 Biochem. Biophys. Res. Commun. 107, 568-574 (1982)
 JOURNAL
 MEDLINE
 REFERENCE
 AUTHORS
 TITLE
 Jans, D.A., Fimmel, A.L., Hatch, L., Gibson, F. and Cox, G.B.
 An additional acidic residue in the membrane portion of the b-subunit of the energy-transducing adenosine triphosphatase of Escherichia coli affects both assembly and function
 Biochem. J. 221, 43-51 (1984)
 JOURNAL
 MEDLINE
 REFERENCE
 AUTHORS
 TITLE
 Walker, J.E., Saraste, M. and Gay, N.J.
 Nucleotide sequence, regulation and structure of ATP-synthase
 Biochim. Biophys. Acta 768, 164-200 (1984)
 JOURNAL
 MEDLINE
 REFERENCE
 AUTHORS
 TITLE
 Gay, N.J.
 Construction and characterization of an Escherichia coli strain with an uncl mutation
 J. Bacteriol. 158, 820-825 (1984)
 JOURNAL
 MEDLINE
 REFERENCE
 AUTHORS
 TITLE
 [9] revises other sequences.
 The atp synthase operon consists of nine genes coding for nine structural proteins of ATPase. It is not known through these

gene
CDS
/db_xref="taxon:666"
/chromosome="1"
/note="biotype: El Tor"
136..2307
/gene="VC2758"
136..2307
/note="VC2758"
/note="similar to GB:M87049 SP:P21177 GB:M59368 GB:M74164
PID:145900; identified by sequence similarity: putative"
/codon_start=1
/transl_table=11
/product="fatty oxidation complex, alpha subunit"
/protein_id="AAE95897.1"
/db_xref="GI:9657359"
/translation="MTQAKTLOVKOLANGIAELSPCAPSVNKLDTLHTLESIDKALD
ATAASSVKGILLSSDKKAFIVGADTTEFLGPAKPEALDEMLPANKLELPGA
PFTLSAKHTLGGGCECYLATDPRIGDPTTSIGPEPLKIMPGCGVRLPRILGA
DSAMETITQKACRAEELKVLGDLIVDSKLDISATITLTOALEKIDWKRRQK
TSALETSLKLEAMMSFTMAKGVAVAGKHYPAPMTSVTIEEARLPRDAALDERKH
FLAKSTEAOALVIGIFLNDQYIKGLAKOSAKAASODTQHAALVAGINGGIAVQSA
FIKALPSTEAOLVIGIFLNDQYIKGLAKOSAKAASODTQHAALVAGINGGIAVQSA
LDVIVEAVENPKYKAAVLSVEGLVOTETILTSTPTINLAKSLKRPNGFAM
QADYVEAVENPKYKAAVLSVEGLVOTETILTSTPTINLAKSLKRPNGFAM
HFPNPNRHPVEIINGEHTSEDTIRVAVYAAKKKSPILVNDCCGPFVNRPLPEYF
AGESTLMRDGANFTEIDKVMERQFPMPCPATILDVAGIDTAHQAQVMAEGPPTMA
KSGREIDALYEAKKFGKNGSGFYQYTDKRGKPKAFSDVLAIALAPVCGAQSFQ
POTLIERMTIPMINEVYLCLEGIILASQAEDMALVYGLGFPFGVGRVRYLDTIGIA
NVYAAEKYADIGALYQVLPQLKNNMAOGTSFYSAQOVSAL"
2333..3496
/gene="VC2759"
2333..3496
/note="VC2759"
/note="similar to GB:M87049 SP:P21151 GB:M59368 GB:M74164
PID:145901; identified by sequence similarity: putative"
/codon_start=1
/transl_table=11
/product="fatty oxidation complex, beta subunit"
/protein_id="AAE95898.1"
/db_xref="GI:9657360"
/translation="MNTVVIVDCLRTPMKSGKGAFFRQRAEDLSAHMKGILARNPO
VNRKTEIDYMGVCOOTLEGGFNVARNAALLAGLPEIGAVTNRLCGSSMOLHDA
RAIMWDAEICLYGVYEHKGRHVPTGVDHPRLSKNVAKAAMGILATAMLKINGI
SRQODDEFARSHARAHAAATLBGRFKEIILPTGSHADGTLFLIDDEVIRPETTYAG
LAEIRVFPDANGTAVAGTSALSDASAMLVASEOKAKALGLTIRARITAMAVAGD
PSIMGYVPVATKALQAGLTMQDMDVLENEAFQAOSLPCKADGLLEMDKYNL
NGCAILGHPLGSGSTRISTTLINLEAKDAKAGLAWGIGGIAITPERB"
/complement(3563..4471)
/gene="VC2760"
/complement(3563..4471)
/note="VC2760"
/note="similar to SP:P37598 GB:U00096 PID:1742732
PID:1742742 PID:1787949; identified by sequence
similarity: putative"
/codon_start=1
/transl_table=11
/product="transcriptional regulator, LysR family"
/protein_id="AAE95899.1"
/db_xref="GI:9657361"
/translation="MMPKSSLEMLDVAALGSTAAEVLHKVPSAISYSVQRIESE
LGVVLRRLPRKVELTPAGELFMLEARQLRQMEITRAQTKRAHGRKTLKTLIDVY
VKIDKMPVEAVYQTFDEALQINMEVENGSEWALAQGRADIVTGAFAVPGGDE
VRMGILDMAFVAVSNPHCVREONLSEGLTSQFLAICLIDTSSVLPKRHEHYKQR
LLPLNYSALIESLSGLGVGMPRHMAQPLLASGOLVEKVLPRDKLLSHCCLVWRKD
NHLTIEMVAVYLGSSDQLYSDMWIOAK"
4600..5829
/gene="VC2761"
4600..5829
/note="VC2761"
/note="similar to GB:U00096 PID:1742733 PID:1742743
PID:1787950; identified by sequence similarity: putative"
/codon_start=1
/transl_table=11
/product="multidrug resistance protein"
/protein_id="AAE95900.1"

gene
CDS
/db_xref="GI:9657362"
/translation="MNPEDGHEKMKVSPFOYLVTLSALSMGFIATDMYLPAFKMETA
FATGEQIALSLTVFLVGMALQDLMLGASDFGFRNTLILALVFTLASAGLASSST
VWOLLILRFIQAICVAPAVIQAAMVIRKYPSSVSOQLPATIMPLVSLAPALPOLGV
LLADYFGMSIFLTLGMLIVYSTGQTNELIAPKETSIMODIKGLKSTYVIGNV
MMFACASAFPAVILGMPETMAOLGQADIDLSFIPOTIAFMVGGYGLKRAVQKYG
EVLKQIALFTLSAMTEFVASQWTLSSITWILAFELIIVANGAMYPIVNRALSSA
KOSPATAGLQNSLQISVSSLSALVMAASQAOVYTGIAIYLCIGGLMLGYVLSNR
LREQFTPDNSRVVSD"
/complement(5891..7252)
/gene="VC2762"
/complement(5891..7252)
/note="VC2762"
/note="similar to SP:P17114 PID:290578 PID:43267
PID:290579 GB:U00096; identified by sequence similarity:
putative"
/codon_start=1
/transl_table=11
/product="UDP-N-acetylglucosamine pyrophosphorylase"
/protein_id="AAE95901.1"
/db_xref="GI:9657363"
/translation="MKFSTVILAAKGRTRMHSNMPKVLHTLAKPMAVKHVIDTCNNLG
AONIHLYVGGGDMOQALANENVNVVLOAQIGTHAVDASPHRODEKTLVYGD
VPLISDTEISLEAQPTDGIALLTVLEDPYGRIVRRGPVVAIVQKDASEOK
LKEVNTGVYVARGRDLKRMIAGLNNNAOGEVYLDVIAAARDEGRAVEAVPSHSI
IEGVNDRIQALAEAFQARQAKKILEGVYMLRDPARDLREGTCCGSDVEVDVVI
IEGVNSIGNNVYVAGSILKDEIDNVTYIRPVSVEGATVGENCTVGPFTLRPAE
LRDVAHGVNFEEMKNAKLGEKSKANHLTYLGADELGVNAGVITTCNDGANKRT
VIDGDVFGSDCLVAPVTIGNCAFTIGACTTTLKVNABESLVTTRAPERKIAQWOPRA
KKK"
/complement(7499..7921)
/gene="VC2763"
/complement(7499..7921)
/note="VC2763"
/note="similar to SP:P12988 PID:48340; identified by
sequence similarity: putative"
/codon_start=1
/transl_table=11
/product="ATP synthase F1, epsilon subunit"
/protein_id="AAE95902.1"
/db_xref="GI:9657364"
/translation="MAITPDLVDSAEKLFGLVETFOVTSSEGLGFIHGHTPLL
TAIKPQVRIYKLGHEHYVYVSGIYVOPGATVLAADAIAGEDLDAAKABEAKR
AEERILQHGDMDFQAASLAKAIAQLRVIELTKRR"
/complement(7940..9343)
/gene="VC2764"
/complement(7940..9343)
/note="VC2764"
/note="similar to GB:X16050 SP:P12986 PID:48339;
identified by sequence similarity: putative"
/codon_start=1
/transl_table=11
/product="ATP synthase F1, beta subunit"
/protein_id="AAE95903.1"
/db_xref="GI:9657365"
/translation="MARGKIVQITGAVVDEPFOSEYPSYDALNVDSKERVLVEVQ
QOLGGGVIRIVMSSDGLRRGMTVONTGAPISVPVGTGLGRIMNVLSAIDBERGIL
GAEEVYSIRHPAPSYEEOSSATELLETGVADILICFPAGKRIELGSGAVGTVM
MELINIAILOSIGISVAGVGERRENDYHEHGEAVNVNVEPELSKAMVYGOVN
EPGNRIRVALITGCTMAEKRRDRGDDVLEIDNRYRTLAGTEVSALGMPSPVVGQ
PTLAEEGVQIERTITTKKSTYSVQAVYVPADDLTPDSATPAHDAIVVNRNIA
AMGYPADIDPDIESTRQDLPLVVGQEHYDARGVQATLQKRELKIDIALILGDELS
ADKQVAVARRKIERFTQPYHVAEVFPGDGVYVPLKETLRFKFGILLAGDYDDIPBOA
FMYCGTIDDAIENAKKL"
/complement(9378..10244)
/gene="VC2765"
/complement(9378..10244)
/note="VC2765"
/note="similar to SP:P12990 PID:48338; identified by
sequence similarity: putative"
/codon_start=1
/transl_table=11
/product="ATP synthase F1, gamma subunit"

```

Db 4475 CCAGAAATATACGAAAGGATGAAATGTAGACACTTCACAAAATATGCGTTCTCAAG 4534
Qy 473 aacgcatgctcttcaagcccttattcaagaacaataagtaagtaagtaagtaagtaag 532
Db 4535 ATGCAATGAAAGATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4594
Qy 533 ccaaaagcaagcttggttcaagcatcatttttaagtgatcgcaagtaaaaaaagtgg 592
Db 4595 CTAAAGCAAACTAGATGATCCGTATCCGTACCTAGAAAGCGTGAAGCTAAACGTGTTG 4654
Qy 593 gcatgatttggtgctccacaagatcggtgcttggtggtggtggtggtggtggtggtggt 652
Db 4655 GTTACATCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4714
Qy 653 aacatgataataatgaataagaaatggaagaagaagaatgttccgttcaatgtagt 712
Db 4715 AAAAAGCCGTTACAGACAAACATGAGAGAGAAAGGTGCTGAGATTGACCTTGCCG 4774
Qy 713 taatcggttctaataatcattcaacttttccaaatcttggaatttaaatttaaccgaag 772
Db 4775 TAGTGGCTCAAAAGCAACAGCTTTCTTTAAACATGTGGCGCAAAAGTTGGCGCACAG 4834
Qy 773 attcaggtattggtgatactccctctgttgagcagtttaattggttcaagtaattca 832
Db 4835 TTTCGTGGCTGGGTGATTAACCAAGCTTAGAAGACTTAATGCTGTTGGCTGATGCG 4894
Qy 833 ttgatgctataaaaaaggggaagtaagtaagtaagtaagtaagtaagtaagtaagta 892
Db 4895 TCAGAAATATACATGAAAGGTGAGCTGAGCTGATACATGATTTCAACAAATTTGTTA 4954
Qy 893 aacagatgctgcaaaaagcagtaattggaataataattcattaccagaattagaatg 952
Db 4955 ATACATATGTTTACGACACCAAGCATGATCAATTTGCTACCTTTGCTTAATGCGAAG 5014
Qy 953 atgaaatgagcgaaagaaacagtttggatataattcaagacctgaltgcgaagat 1012
Db 5015 AAGAGAT---GCAGCGGTGACACTCATGAGGACATCATATGATGAGCTGAGCCAAAC 5071
Qy 1013 tatttgataattatttgggttcttatttagaatctcagtttcaagcagcagttgaa 1072
Db 5072 TACTGATTCAGCTACTAGTACGCTTACGAGTCTGAGTATACAAAGGTGTGTGTGA 5131
Qy 1073 acctggtcttgagcaagcgcgtcgaatggtgcgcataagcaagcaagcaagcaag 1132
Db 5132 ACCTGCTCTGTGAACAGCGGCTCGAATATTCGATTCGATGAAGGTGCAAGGATTAAG 5191
Qy 1133 gtaacttaataatgagttacagttatgctatacaaaagctcgtcaagaagtattaca 1192
Db 5192 CCAACTGATGACGATTTGGAACCTGTGTACAAACAAAGCCGTCAGGCTGCGATCAC 5251
Qy 1193 atgaatataatgaatgttgcgcgtgcagcagcaattaa-----c 1234
Db 5252 AAGAAATATTCGGAACATCGTTGTGTGTGAGCTGCGGTTAAAGCTTAGTAAACGA 5311
Qy 1235 aataagagatcgtaatggaactggaataattgtaacaatcatcggtgagttatga 1294
Db 5312 GTTTAAGAGATTAAAGATATGCTTACAGAGTATGATGATGATGATGATGATGATG 5371
Qy 1295 cgttgaaatccacaagatgacgtacaaagtataatgacgttcaatgt----- 1345
Db 5372 CGTAGAGTTCCACAGACAGATGATGATGATGATGATGATGATGATGATGATGATG 5431
Qy 1346 tgaagaagtttagtactcgtgaagttcaacaacaataagttggtggtggtggtggt 1405
Db 5432 AAAAGAGCCTTATGTTCTTGAAGTTCAACACAGCTGAGCGGCGGTGATGTTGAT 5491
Qy 1406 cgcgaatggagatcatcgtgagtttaaaacgcggttaagcgtaacaataagataacc 1465
Db 5492 CTTAATGGTACGCTCTGATGCTTACGCTGTGAGAGTTGAAGTGTGAACCTGCGCTCC 5551
Qy 1466 aatttcgttcagtgagcaaaaacatttggttcgtatcatatgaagtttgggtgaac 1525

```

```

Db 5552 AATTTCAGTACAGTCGCTACTTAAACCTTAGTCTGATATGAAAGCTCTAGTGACGC 5611
Qy 1526 aatcagatgagcaaggtgaaatcggtgcagaagaagtgtgtattcaccgtgcgcac 1585
Db 5612 GATTGATGAGCGGTGTGAAGTGGCGGAGAGGTTTACTTATCTACACCGACGCCACC 5671
Qy 1586 aagtttgaagaacaactcaactgaactgaactttagaagaaggaatlaaagttacga 1645
Db 5672 AACCTAGCAAGAAACATCAAAAGCATTCGACTTCTAGAAAGCGGGTTAAAGTATTCGA 5731
Qy 1646 cttaagtttgcgttgcgaagaagggtaagaagtttattcctggtggtggtggtggt 1705
Db 5732 CCTAATTTGCTCCATTCGCTAAGGCTGTAATAATCGCTTATTCGCTGAGCTGAG 5791
Qy 1706 taaacgtcaatataatgaatlaacgttaacgtcaacgtcaacgtcaacgtcaac 1765
Db 5792 TAAGACCGTTAATATATGAGAACTTATCAACATCGACCTACACACTCAGGCTTATC 5851
Qy 1766 tgccttcgagggtagtagtagtagtagtagtagtagtagtagtagtagtagtagtag 1825
Db 5852 CATTATTCGCGGTGTGTGAGCGTACTGCTGAAGGATTAACGATTTCTACAGAGATGCA 5911
Qy 1826 agacttcaagttataga-----taagttctcttctgtttagg 1864
Db 5912 GGAAGCGCGCTGTGAAGCTTGAAGCTTGAAGATGCAAAAGTACGATGCTTTACGG 5971
Qy 1865 tcaatgaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 1924
Db 5972 TCAGATGAACGAGCACACAGGTAACCGTCTACGTTGTGCACTGACGTGCTTAACA 6031
Qy 1925 ggaanaattcgtgataaggtcgtatgtcttattcttcttcttcttcttcttcttct 1972
Db 6032 AGAAGCTTTCCTGAGCAAGGTCGTACGATCACTGATCATTTGATTA 6079

RESULT 13
AE004342/c
LOCUS
DEFINITION
Vibrio cholerae chromosome I, section 250 of 251 of the complete
chromosome.
ACCESSION
AE004342 AE003852
VERSION
AE004342.1 GI:9657358
KEYWORDS
SOURCE
ORGANISM
Vibrio cholerae.
Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
REFERENCE
1 (bases 1 to 12928)
Heidelberg, J.F., Eisen, J.A., Nelson, W.C., Clayton, R.A., Gwinn, M.L.,
Dodson, R.J., Haft, D.H., Hickey, E.R., Peterson, J.D., Umayam, L.,
Gill, S.R., Nelson, K.E., Read, T.D., Teitelin, H., Richardson, D.,
Ermolaeva, M.D., Yamathavan, J., Bass, S., Qin, H., Dragol, I.,
Sellers, P., McDonald, L., Utterback, T., Fleischmann, R.D.,
Nierman, W.C., White, O., Salzberg, S.L., Smith, H.O., Colwell, R.R.,
Mekalanos, J.J., Venter, J.C. and Fraser, C.M.
Direct Submision
Submitted (14-JUN-2000) The Institute for Genomic Research, 9712
Medical Center Dr. Rockville, MD 20850, USA
FEATURES
source
1..12928
/organism="Vibrio cholerae"
/strain="NI6961"
/serotype="O1"

```

QY	1827	gactctaagctataga-----taagtgctcttgittatgct	1865
Db	8758	GAACGGGGCGTTGTTAAACGTTGAACAACAGACTGCGAAAGTAGCGATGTTTACGGT	8699
QY	1866	caatgaacagagcaccacaggttaaccgtttcagtgatgcatgaacagcttaactatgacg	1925
Db	8698	CAGATGAACGAGCCACCAAGCAACCGTCTGCGTGTACACTGACTGACTGCTGCAGCATATGGC	8699
QY	1926	gaaaattccgtgatgaagtcgctgatgctcttattcttcgttgataa	1972
Db	8638	GAAGAATTCCTGATGAAGCGCTGAGACTGCTGTTATATGCACAA	8592
RESULT	14		
LOCUS	AP001118		
DEFINITION	AP001118 347550 bp DNA BCT 07-SEP-2000		
ACCESSION	Buchnera sp. APS genomic DNA, complete sequence, segment 1/2.		
VERSION	AP001118.1 GI:10038693		
KEYWORDS	ferredoxin II; GTP-binding protein; membrane protein; kinase.		
SOURCE	Buchnera sp. APS (strain:APS, isolate:Tokyo1998, specific_host:Acyrthosiphon pisum) DNA.		
ORGANISM	Buchnera sp. APS		
REFERENCE	Bacteria; Proteobacteria; gamma subdivision; Buchnera.		
AUTHORS	1 (sites) Shigenobu,S., Watanabe,H., Hattori,M., Sakaki,Y. and Ishikawa,H.		
TITLE	Genome sequence of the endocellular bacterial symbiont of aphids		
JOURNAL	Buchnera sp. APS		
MEDLINE	Nature 407 (6800), 81-86 (2000)		
AUTHORS	2 (bases 1 to 347550) Shigenobu,S., Hattori,M., Watanabe,H., Toyoda,A., Yada,T., Sakaki,Y. and Ishikawa,H.		
TITLE	Direct Submission		
JOURNAL	Submitted (24-JAN-2000) to the DDBJ/EMBL/GenBank databases. Shuji		
MEDLINE	Shigenobu, University of Tokyo, Department of Biological Sciences;		
AUTHORS	7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan		
COMMENT	(E-mail:shige@sg.riken.go.jp, URL:http://buchnera.gsc.riken.go.jp, Tel:81-3-5800-3553, Fax:81-3-5800-3553)		
FEATURES	Collaboration Information: RIKEN, Genomic Sciences Center (GSC)		
SOURCE	Kitasato Univ., 1-15-1 Kitasato, Sagamihara, Kanagawa 228-8555, Japan		
FEATURES	Tel:81-42-778-9923 FAX:81-42-778-9924		
SOURCE	AP000398: Submitted(23-Aug-1999).		
LOCUS	Location/Qualifiers		
LOCUS	1. 347550		
LOCUS	/organism="Buchnera sp. APS"		
LOCUS	/isolate="Tokyo1998"		
LOCUS	/specific_host="Acyrthosiphon pisum"		
LOCUS	/strain="APS"		
LOCUS	/db_xref="taxon:107806"		
LOCUS	/note="primary endosymbiont of Acyrthosiphon pisum"		
LOCUS	197. 2083		
LOCUS	/gene="gida"		
LOCUS	/gene="gida"		
LOCUS	/note="BU001"		
LOCUS	/codon_start=1		
LOCUS	/transl_table=1		
LOCUS	/product="glucose inhibited division protein A"		
LOCUS	/protein_id="BAB12729.1"		
LOCUS	/db_xref="GI:10038694"		
LOCUS	/translation="MFNRNDVIVYAGAGHGTAAAMSSMGKTLILNOKISDGA		
LOCUS	LCSPAIGIGIKSHLVKEDIDLGGMAKADISIGIRILSSGPAVRSTRADADIT		
LOCUS	LYHTHTKILKRNKLLILEAEVDLLEKNSVVGVLQNEINFSVSLAACTFLIG		
LOCUS	GHIIIGKTSYAGRIIGKRSALDVSRLREKSLRVNRLKGTPPRIDINTVFNLLIQ		
LOCUS	NSDPVIVPEFMGNVSHHPKOIPCYLHTNKEKTEIRKLNLDKSPITVGFKGLGPRY		
LOCUS	CSPIEDKIVPRPDKRSQVLEFEGSLSSIKVYNGISTSLPIEVOEDIVASIGLEKS		
LOCUS	KIIRGVAIETDFDPKDLNLTESKLKLGKLFPGQINGTGYEAAASQGLAGLNAA		
LOCUS	LSSTKSWEPFRDQATVGLVLDITLQTQGEPEYRMFTSRAEVLSLRDNADLRLE		
LOCUS	IGRKIGLVNDSRWIRYNOKVNIQTENMRLLKNNISPISPADILIKLNIINLKEIS		

gene
2278. .3102
/gene="atpb"
2278. .3102
/gene="atpb"
/EC_number="3.6.1.34"
/note="B0002"
/codon_start=1
/transl_table=11
/product="ATP synthase A chain"
/protein_id="BAB12730.1"
/db_xref="GI:10038695"
/translation="MILEKISDPQKYISHLSHQIDLRSEKIIQPCALSDSYTYV
DSMPFSVLGSEFLISFYMWCKITQGIQPKLQTAIELIEFVNLNYSWYQKQNALI
ALSLTVAIWEIMLMDLVPIDEPFSEKEVPELPMRIVPSPADINITLSIGVEF
LILFTYVAKIGKVGFKELIILQPFNHPVESIFNFILDEVSLVSPKISLIGRLGNMA
GEMIFITLALGPMWTCQFLVNPVPAIFHILISQALFIPWLTIVYLSMASQSHKD"
3139. .3378
/gene="atpe"
3139. .3378
/gene="atpe"
/EC_number="3.6.1.34"
/note="B0003"
/codon_start=1
/transl_table=11
/product="ATP synthase C chain"
/protein_id="BAB12731.1"
/db_xref="GI:10038696"
/translation="MENINVDMLTYAVAWYGLASIGALIGILGSKFLEGARQPD
LVPILRTQFFVWGLVDALPMIAGSLGYMLFAIS"
3497. .3982
/gene="atpf"
3497. .3982
/gene="atpf"
/EC_number="3.6.1.34"
/note="B0004"
/codon_start=1
/transl_table=11
/product="ATP synthase B chain"
/protein_id="BAB12732.1"
/db_xref="GI:10038697"
/translation="MNLNRTIIGQALISVLYVWRCMKYIMPTIILAIETROKEIKESL
TNAKRKDELTYLEKRIHONITIDAKQKSNLNSANKQVSILEDARQOALESKKSLI
LNTQSEINAIATHARKNLKEVVDLSISMAEKRIKKNSKIDNOELDELVTSLSQVK
N"
3982. .4515
/gene="atph"
3982. .4515
/gene="atph"
/EC_number="3.6.1.34"
/note="B0005"
/codon_start=1
/transl_table=11
/product="ATP synthase delta chain"
/protein_id="BAB12733.1"
/db_xref="GI:10038698"
/translation="MSVADTIARPYAQAIFELAIENNTIERKWNILIFIKTIASHKKR
KNLSGISPSKYLISLIFITGNIIDENAKNLIKLSNORSPNLIANNIEFVFKLEAG
YKIIITVQKSAFSLKENLINKIRKYLERFIAKKIITIKYDPTNLNMIVKVNNTIE
DLSAQNHKLQSLDSLNF"
4530. .6068
/gene="atpa"
4530. .6068
/gene="atpa"
/EC_number="3.6.1.34"
/note="B0006"
/codon_start=1
/transl_table=11
/product="ATP synthase alpha chain"
/protein_id="BAB12734.1"
/db_xref="GI:10038699"

Query Match 38.8%; Score 764.8; DB 2; Length 347550;
 est Local Similarity 64.2%; Pred. No. 7.7e-140;

	Matches	1267:	Conservative	0:	Mismatches	657:	Indels	48:	Gaps	6:
QY	43	cacacgtaagcaat	tatcatcaat	tggtcaaa	aaagtaac	tgaaatatt	tgaaacaa	accata	102	
Db	5777	CACGTGAAAGCA	ATTAAATCAAT	GTCTCAAAA	AAATATTCAG	CAATATTAA	CAAAAA	CAATA	5836	
QY	103	ctctccgttattc	tgtagcaaca	aaagcaatt	tggtattt	gtgaagat	ggtt	ggttact	162	
Db	5837	TTCCACCTATT	TTCTTAATAG	CGGACAG	CGCTTATAC	TTTGTGTG	GAGAAAT	AAATTTTCT	5896	
QY	163	agaagaagtgag	acttagatc	gcatcgtt	gcatcttgat	taacgaac	ctttagat	gagta	222	
Db	5897	TGATGATATT	TTCTATAG	ATTAATAC	CTTAATTTG	AAAAAGAA	ATTTATTTAT	TATGCTCA	5956	
QY	223	ccataactatg	ctgattt	ttatgcgt	gagtt	taaccca	ctctg	gcaat	282	
Db	5957	TAAATTATTAT	TTTGGTTTAA	TGAGAAAT	TTATTAAC	AGCGCAT	TTTTTAATTT	TTGTTAT	6016	
QY	283	taaaagatcatt	aaagcatt	tttgtagt	cttcaag	caaacag	tgcgt	tgtaagtt	342	
Db	6017	AAAAATATA	TTTATTATTA	TATTAATTA	CCGAGTTT	AAAAAAAT	TCATTTT	AAAAAAATPA	6076	
QY	343	caacttaaatg	agagacaaa	---tg	cgagtg	ctaaag	atagata	agaacaa	399	
Db	6077	CTAATTAAGT	TTAAAAAGAG	AAAAATAG	TGACAC	ACACAAAA	AAAAATTTAAAA	ACAAAAATTCGG	6136	
QY	400	agtgtaaaag	tacacaa	aaatctaa	aagcgat	gtgaaat	tggtgtc	gtcctcg	459	
Db	6137	AGTGTTACCA	ATTAATTA	AAATTCAT	TAAGCG	ATGAGAA	TGGTGGC	AGTGCTFAAATG	6196	
QY	460	cgtaaaagc	gaagacg	catgctctt	cttcacg	cccttat	cagaac	aatagta	519	
Db	6197	ACAAAAAC	GGAAGAA	GAATGCG	ATCAGG	TCACCAT	ATTCTG	ATTAATTA	TGAAAAAGTT	6256
QY	520	attagccagctt	ccaagca	cgattg	gttaca	agcatcc	attt	agtgatc	579	
Db	6257	ATTGATCAT	GTGACAC	AGGAATTTAG	ATATTAAC	ATAGTTACT	GTTCG	ACAGAAAGAAA	6316	
QY	580	gtaaaaaag	tgggatgat	tgltgt	gtccaca	agatc	gtgtc	gttgtgt	639	
Db	6317	ACTAACCG	ATTGGCAT	GTATTAAT	TATCTACT	GTATG	ATGAG	ATATTAAT	6376	
QY	640	gtgaacttg	tttaaa	actcgtat	taata	tgaaat	tgaaagaa	tgaaagaa	699	
Db	6377	ACTAATCT	TTTTTAA	CAAGATATTA	TATTAAC	ACGAATTTG	CTPAAG	ATAATTTCCG	6436	
QY	700	gtccaatgag	ttatcggt	ctcaatc	atcaactt	ccaactt	ccaactt	gtgga	759	
Db	6437	TGTGATTTA	TATTTATTTG	TTTAA	AAAGTTTAT	CTGATTTCA	AACTTTG	TGTGGAAGTAT	6496	
QY	760	atttaacc	caagattc	ggtat	gtgtga	tactcc	ctctgt	gtgag	819	
Db	6497	ATTCTTGGA	ATGAAGCA	CTACTAG	GAGAAAT	TCCCAAT	TAGAG	GAATTAATTA	6556	
QY	820	gtaacttcta	gatgat	gtctata	aaaa					


```

Db 231 CTGGAATATTAAGACACCTTACCTGGAAGACCGGAGCTTAAACGGCTGGGCTACCTGGTG 290
Qy 604 ggtccacagatcgtggtcttctgtgtggtggttaaacgtgaactgtttaaactgtatla 663
Db 291 GTGTCCACCGACCGCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 350
Qy 664 aatgaatgaagaatgaagaagaagaagaatgttccgttcaattgaattgaattcgttct 723
Db 351 GCGGAATGAAGAACCTGAGACCGGACCAAGGCGTCAATGGACCTCCCAATGATGGCTG 410
Qy 724 aaatcatcaactatcttcccaatcttctgggaattaaatlaaaccaagattcaagtatt 783
Db 411 AAGAGCGTCTGCTGTCTTCAACTCCGCGGCGGCAATGTTGTGCCAGTCCACCGCATG 470
Qy 784 ggtgatactccctctgttgaagcaatgaattgttcagtlcaatcttctatgaattgtat 843
Db 471 GGGGATTAACCTTCCCTGTCGCAACTGATCGGTAAAGTGTATGTCAGGCGCTAC 530
Qy 844 aaaaaaggggaagtgaattgtgtatattgaattataaacaattatlaacagatgtcg 903
Db 531 GACGAAGGCGCTGTGACACAGCTTACATTTACATGACGCAACAAATTTAATTAACACATGCT 590
Qy 904 caaagccagatttggaaanaattcaattcattaccagaattagataatgaattagggc 963
Db 591 CAGGTTCCGACCATGACGACGCTGCTGCGCTTACCGGCTACGATGATGATGATCT---- 646
Qy 964 gaaagaaaaaagatttggatataatlaacgaactgaatgcgaagaattatlaagataat 1023
Db 647 --GAAACATTAATCCGAGGATTAACCTGTACGAACCCGATCCGAAGCGCTTGGTGATAC 704
Qy 1024 ttattgttctgtatttagaatcgaattcaaggttatacagcgcagttgaaacttctct 1083
Db 705 CTGCTGCTGCTGTATGTCGAATCTCAGGTTTACAGGGCGGTGTAACAACTGGCCACG 764
Qy 1084 gagcaagccgcctgaattgttcgcaatgaagcagcaagataacgcaagtaactaatt 1143
Db 765 GAGCAGGCGCGCGGCTATGCTGTGCGATGAAGCCGCGACCGACATGCGCGGACCTGATT 824
Qy 1144 aatgaattcaagttagttcctatacaaaagctgcgaagcaagattatacaaatgaat 1203
Db 825 AAGAGACTCAGTGTGTATATCAAAAGCTGTCAGGCGACGATTAATCAGGAATCAACC 884
Qy 1204 gaaattgttgcggtgagcagcaatttaacaa-----atagaagatcgttaattg 1254
Db 885 GAGATCGTCTCGGGGCGCGCGGCTTAAACAGGTTATTCGTAGAGGATTTAAGATGG 944
Qy 1255 caacttggaaaaatttgaacaatcatcgtgtggttattgaattccacaaagatg 1314
Db 945 CTACTGGAAGAATGTGTCCAGATATCGGCGCGGTGATGACGTGCAATTCCTCAAGATG 1004
Qy 1315 cagtaaccaaaagtataatgactt-----aaatgttgaacaggtttagtactg 1365
Db 1005 CCGTACCGCGCGCTGTACGATGCTTGAAGTCAAAATGATGAGCTGTGCTGTGG 1064
Qy 1366 aagttcaacaacaattaggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1425
Db 1065 AAGTTCAGACGAGCTCGGCGGCGGTATCGTACGATCCGCAATGGGTTCTCTCCGAGC 1124
Qy 1426 gattaaaaagcgtttaaagcgtlaacaacaatacgaatacccaattctgttccagttggaa 1485
Db 1125 GTCTGCGTCCGCGTCTGATGTAAAGAACCTCGAACCCTGATGAAGTCCCGGTAGGTA 1184
Qy 1486 cgaaaacatttgggtcgtatcatgaacgtatgtgtgaaccaaactgaagcaagttgaa 1545
Db 1185 AAGCGACTCTGGGCGGTATCATGAAGCTGCTGGTGAACCGGTGACATGAAGAGCGAGA 1244
Qy 1546 tgggtgcgaagaagaatgtgttatttccacgtgcgcaacaaagtataagaacaactta 1605
Db 1245 TCGGTGAAGAAAGAGGCTGTGGGCGATTACCGCGCAGCACCTTCTCAAGAAAGCTGTCAA 1304
Qy 1606 acagtaactgaactttagaagaacgggaatgaagttatcgaacttagtttccgtttgca 1665

```

```

Db 1305 ACTTCAGGAAGTGTGGAACCCGCTATCAAGATTAACGACTGATGTGCTGCTGCTA 1364
Qy 1666 aagggtgaagaatgaattatcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1725
Db 1365 AGCGCGGTAAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1424
Qy 1726 aattatcgttaacatgcaatttgaacgaactgaagttactgttctgttgcgggtgagtg 1785
Db 1425 ACCTATTCGTGAACATCCGATCGAGACACTCCGCTTACTGTGTGTGTGTGTGTGTGTGT 1484
Qy 1786 agcgtacggtgaagtaacgaacttctatcatgaatgaagaactcaacgtatlaagata 1845
Db 1485 AACGTACGTGTAGGCTTAACGACTTCTTAACGAAGATGACGATCCCAAGTTATCGACA 1544
Qy 1846 aagttctctgttattgttcaaatgaacgagccacagtaacggttlaagttgtgcat 1905
Db 1545 AAGTATCCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1604
Qy 1906 taacagcttaactatgtgcggaanaattcgtgtatgaaggtcgtgtatcttattcttg 1965
Db 1605 TGACCGGTCTGACCATGAGCTGTGAATAATCCGTGACGAAGGTGTGACGTGTGCTGTG 1664
Qy 1966 ttgataa 1972
Db 1665 TTGACAA 1671

```

Search completed: October 28, 2001, 16:35:31
Job time: 5810 sec



GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 28, 2001, 15:00:31 ; Search time 141.96 Seconds
(Without alignments)
8722.330 Million cell updates/sec

Title: US-09-545-199c-3
1972
Perfect score: 1 agcggcgccattggtcagc.....gtctattcttcgtgataa 1972
Sequence:

Scoring table: IDENTITY-NUC
Gapop 10.0 , Gapext 1.0

Searched: 730101 seqs, 313950809 residues
Total number of hits satisfying chosen parameters: 1460202

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N.Geneseq_0601.*
1: /SIDSB/gcgdata/geneseq/geneseqn/NA1980.DAT.*
2: /SIDSB/gcgdata/geneseq/geneseqn/NA1981.DAT.*
3: /SIDSB/gcgdata/geneseq/geneseqn/NA1982.DAT.*
4: /SIDSB/gcgdata/geneseq/geneseqn/NA1983.DAT.*
5: /SIDSB/gcgdata/geneseq/geneseqn/NA1984.DAT.*
6: /SIDSB/gcgdata/geneseq/geneseqn/NA1985.DAT.*
7: /SIDSB/gcgdata/geneseq/geneseqn/NA1986.DAT.*
8: /SIDSB/gcgdata/geneseq/geneseqn/NA1987.DAT.*
9: /SIDSB/gcgdata/geneseq/geneseqn/NA1988.DAT.*
10: /SIDSB/gcgdata/geneseq/geneseqn/NA1989.DAT.*
11: /SIDSB/gcgdata/geneseq/geneseqn/NA1990.DAT.*
12: /SIDSB/gcgdata/geneseq/geneseqn/NA1991.DAT.*
13: /SIDSB/gcgdata/geneseq/geneseqn/NA1992.DAT.*
14: /SIDSB/gcgdata/geneseq/geneseqn/NA1993.DAT.*
15: /SIDSB/gcgdata/geneseq/geneseqn/NA1994.DAT.*
16: /SIDSB/gcgdata/geneseq/geneseqn/NA1995.DAT.*
17: /SIDSB/gcgdata/geneseq/geneseqn/NA1996.DAT.*
18: /SIDSB/gcgdata/geneseq/geneseqn/NA1997.DAT.*
19: /SIDSB/gcgdata/geneseq/geneseqn/NA1998.DAT.*
20: /SIDSB/gcgdata/geneseq/geneseqn/NA1999.DAT.*
21: /SIDSB/gcgdata/geneseq/geneseqn/NA2000.DAT.*
22: /SIDSB/gcgdata/geneseq/geneseqn/NA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1972	100.0	1972	21	AAC79582
2	603.6	30.6	46593	21	AAAB81456
3	603.6	30.6	349980	21	AAAF21612
4	603.6	30.6	837096	21	AAAB81489
5	498.8	25.3	867	18	AAC79654
6	402.4	20.4	6309	18	AAV74791
7	335.2	17.0	7750	20	AAI13153
8	333.2	16.9	94750	22	AAAF28551
9	297	15.1	580073	18	AAAF58840
10	286.6	14.5	531	21	AAC79635
11	281.2	14.3	10399	19	AAV52293

12	261.2	13.2	4815	19	AAV29571	L. lactis soluble
13	250.8	12.7	4001	20	AAV29857	Thermotoga neopolis
14	215.4	10.9	1468	19	AAI14416	H. pylori GHPD 663
15	210	10.6	8008	20	AAE32024	Human MTH1 relate
16	210	10.6	8009	21	AAE90081	X99599 cDNA clone.
17	196.8	10.0	1812	21	AAE07479	Fusarium venenatum
18	189.6	9.6	1683	17	AAI38409	Mitochondrial ATP
19	189.6	9.6	1683	17	AAI38404	ATP-synthase Fl-AT
20	183.6	9.3	1572	22	AAV71740	Corynebacterium g1
21	182.6	9.3	1904	21	AAC37463	Human cancer assoc
22	182.4	9.2	1452	16	AAE03954	Arabidopsis thaliana
23	182	9.2	1407	19	AAV43719	Brevibacterium fla
24	181	9.2	1807	19	AAV43719	Cancer associated
25	179.2	9.1	1931	21	AAE33238	Arabidopsis thaliana
26	175	8.9	408	21	AAE37574	Human ORF1279
27	172.6	8.8	2207	19	AAV29572	L. lactis soluble
28	163.8	8.3	2161	11	AAE29573	S. thermophilus Fl
29	159.8	8.1	134525	11	AAE04525	Total base sequenc
30	154.4	7.8	1011	18	AAE84081	DNA encoding a S.
31	151.4	7.7	936	22	AAE58252	Oligonucleotide D1
32	151.4	7.7	936	22	AAE58254	Oligonucleotide D1
33	151.4	7.7	936	22	AAE58257	Oligonucleotide D1
34	151.4	7.7	936	22	AAE58259	Oligonucleotide D2
35	151.4	7.7	936	22	AAE58262	Oligonucleotide D2
36	151.4	7.7	936	22	AAE58255	Oligonucleotide D1
37	151	7.7	936	22	AAE58252	Oligonucleotide D1
38	151	7.7	936	22	AAE58254	Oligonucleotide D1
39	151	7.7	936	22	AAE58257	Oligonucleotide D1
40	151	7.7	936	22	AAE58259	Oligonucleotide D2
41	151	7.7	936	22	AAE58262	Oligonucleotide D2
42	146.2	7.4	1929	14	AAE043748	Oligonucleotide D1
43	146.2	7.4	1929	14	AAE043748	Mitochondria ATPas
44	136	6.9	739	22	AAE71742	Corynebacterium g1
45	109	5.5	688	21	AAE14649	Apetegillus oryzae

ALIGNMENTS

RESULT 1	
AAE79582	standard; DNA; 1972 BP.
ID	AAC79582
XX	
AC	AAC79582;
XX	
DT	08-FEB-2001 (first entry)
XX	
DE	Virulence gene #2.
XX	
KW	Virulence gene: antibacterial; vaccine; bacterial infection;
XX	septicemia; bronchopneumonia; rhinitis; wound infection; ss.
OS	Pasteurella multocida.
XX	
PN	WO200061724-A2.
XX	
PD	19-OCT-2000.
XX	
PF	06-APR-2000; 2000WO-US09218.
XX	
PR	09-APR-1999; 99US-0128689.
XX	
PR	10-SEP-1999; 99US-0153453.
XX	
PA	(PHAA) PHARMACIA & UPJOHN INC.
XX	
PI	Lowery DE, Fuller TE, Kennedy MJ;
XX	
DR	WPI: 2000-647422/62.
XX	
DR	P-PSDB; AAB44522.
XX	
PT	Attenuated Pasteurellaceae bacteria comprising mutations in virulence genes, useful as a live attenuated vaccine against bacterial infections

xx Claim 1: Pages 67-68; 322pp; English.

ps The family Pasteurellaceae encompasses several pathogens that infect a
cc wide variety of animals. The present invention relates to virulence genes
cc from Pasteurellaceae. The present sequence is one such virulence gene.
cc The present sequence may be mutated in order to produce an inactive gene.
cc The inactive virulence gene may in turn be used to produce a vaccine,
cc which is useful for treating bacterial infections such as septicemias,
cc bronchopneumonias, rhinitis and wound infections.

xx Sequence 1972 BP; 652 A; 312 C; 438 G; 570 T; 0 other:

Query Match 100.0%; Score 1972; DB 21; Length 1972;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1972; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 agcggcattgctcagtcggtcggatctctatgtagtcacgtaagcaattatc 60
    |||
DB 1 agcggcattgctcagtcggtcggatctctatgtagtcacgtaagcaattatc 60
    |||

QY 61 acatggtcaaaaagtaactgaattatgaaacaaacaaactatctccgttatctgagc 120
    |||
DB 61 acatggtcaaaaagtaactgaattatgaaacaaacaaactatctccgttatctgagc 120
    |||

QY 121 acacaagcattggttatttgcaatgagttggttacttgaagaagtggaattaga 180
    |||
DB 121 acacaagcattggttatttgcaatgagttggttacttgaagaagtggaattaga 180
    |||

QY 181 tcgatgtgtcatttgaatcagacacttttagatgtagtactaaccataactatgctgatt 240
    |||
DB 181 tcgatgtgtcatttgaatcagacacttttagatgtagtactaaccataactatgctgatt 240
    |||

QY 241 tatgctgtgattacaaccaatctggcaattacaatgataaattaaagtcattaaagg 300
    |||
DB 241 tatgctgtgattacaaccaatctggcaattacaatgataaattaaagtcattaaagg 300
    |||

QY 301 catttgatagcttcaaacgaacagtcggtgaagttaacacttaaaatggaagagac 360
    |||
DB 301 catttgatagcttcaaacgaacagtcggtgaagttaacacttaaaatggaagagac 360
    |||

QY 361 aaaaatgacggtgtcctaaagataagaacaaacaaatcgcgagtgtaaaagtacacaaa 420
    |||
DB 361 aaaaatgacggtgtcctaaagataagaacaaacaaatcgcgagtgtaaaagtacacaaa 420
    |||

QY 421 attactaaagcgtatgaaatggttgcgtcgcgaagaatgcctaaacgcaagaacgcatg 480
    |||
DB 421 attactaaagcgtatgaaatggttgcgtcgcgaagaatgcctaaacgcaagaacgcatg 480
    |||

QY 481 tctcttcaacgaccttattcgaacaaatacgttaacgttgaatgcaacagttcccaagca 540
    |||
DB 481 tctcttcaacgaccttattcgaacaaatacgttaacgttgaatgcaacagttcccaagca 540
    |||

QY 541 aagatggttacaagcatcattttagtggatcgcgagaagtaaaaaagtggcgcatgatt 600
    |||
DB 541 aagatggttacaagcatcattttagtggatcgcgagaagtaaaaaagtggcgcatgatt 600
    |||

QY 601 gttggtccacaagatcgtggtccttgggtggtggttaaacggaactgttttaaaactta 660
    |||
DB 601 gttggtccacaagatcgtggtccttgggtggtggttaaacggaactgttttaaaactta 660
    |||

QY 661 ttaaatgaaatgaaagatggaagaaagaaatgttccggttcaatttgaatttaactcgt 720
    |||
DB 661 ttaaatgaaatgaaagatggaagaaagaaatgttccggttcaatttgaatttaactcgt 720
    |||

QY 721 tctaaatctatcaacttttccaactcttgggaattaaatttaaccagaatcaggt 780
    |||
DB 721 tctaaatctatcaacttttccaactcttgggaattaaatttaaccagaatcaggt 780
    |||

QY 781 attggtgatactccctcgttggagagtaattggttcaagtaacttcatgattgattcct 840
    |||
DB 781 attggtgatactccctcgttggagagtaattggttcaagtaacttcatgattgattcct 840
    |||
```

```

QY 841 tataaaaagggaagtagatgttgtgtattttagtttaatacaaatatttaacacgacg 900
    |||
DB 841 tataaaaagggaagtagatgttgtgtattttagtttaatacaaatatttaacacgacg 900
    |||

QY 901 tcgcaaaagccgatttggaaaaaatattccattacccagaaattagaataatgataat 960
    |||
DB 901 tcgcaaaagccgatttggaaaaaatattccattacccagaaattagaataatgataat 960
    |||

QY 961 ggcgaagaagaacaaatgttggattatattacgaacctgtagcgaagtattattagatt 1020
    |||
DB 961 ggcgaagaagaacaaatgttggattatattacgaacctgtagcgaagtattattagatt 1020
    |||

QY 1021 aattattggttcgttatttagaatctcaggtttatcaaacgacgttgaaaacctgtc 1080
    |||
DB 1021 aattattggttcgttatttagaatctcaggtttatcaaacgacgttgaaaacctgtc 1080
    |||

QY 1081 tctgagcaagccgctcgaatggtcgcataaaagcagcaacagataacggaattactta 1140
    |||
DB 1081 tctgagcaagccgctcgaatggtcgcataaaagcagcaacagataacggaattactta 1140
    |||

QY 1141 attaatgattacagttagttactataacaaagctcgtcaagcaagttatacaattgatta 1200
    |||
DB 1141 attaatgattacagttagttactataacaaagctcgtcaagcaagttatacaattgatta 1200
    |||

QY 1201 aatgaaatctgtccggtgcagcagcaaatttaacaaatagaggtcgttaattgcaactg 1260
    |||
DB 1201 aatgaaatctgtccggtgcagcagcaaatttaacaaatagaggtcgttaattgcaactg 1260
    |||

QY 1261 gaaaatctgtacaaatcatcagtcggttattgaacgttgaattccccaagaatgacgttac 1320
    |||
DB 1261 gaaaatctgtacaaatcatcagtcggttattgaacgttgaattccccaagaatgacgttac 1320
    |||

QY 1321 caaaagataatgacgttcaaatgttgaacaggttagtactgaaagttcaacaacaat 1380
    |||
DB 1321 caaaagataatgacgttcaaatgttgaacaggttagtactgaaagttcaacaacaat 1380
    |||

QY 1381 taagtgtgtgtgtagttcgtgtatcgcaatgggatactgtaggttttaaaacggtt 1440
    |||
DB 1381 taagtgtgtgtgtagttcgtgtatcgcaatgggatactgtaggttttaaaacggtt 1440
    |||

QY 1441 taagcgttaacaatacgaataaaccaaatttctgtccagtggtggaacgaacaaactggttc 1500
    |||
DB 1441 taagcgttaacaatacgaataaaccaaatttctgtccagtggtggaacgaacaaactggttc 1500
    |||

QY 1501 gttcatgaaacgtatctgtgtggaaccaaactcgatggaacaggttgaatctggtggaagaaga 1560
    |||
DB 1501 gttcatgaaacgtatctgtgtggaaccaaactcgatggaacaggttgaatctggtggaagaaga 1560
    |||

QY 1561 attggttattcacacgtggtggtccacaaagttaagaagaacaaactcaacagttactgaacttt 1620
    |||
DB 1561 attggttattcacacgtggtggtccacaaagttaagaagaacaaactcaacagttactgaacttt 1620
    |||

QY 1621 tagaacaaggaattaaagttatcgaacttagtttgcggttgcgaaggggttaagtag 1680
    |||
DB 1621 tagaacaaggaattaaagttatcgaacttagtttgcggttgcgaaggggttaagtag 1680
    |||

QY 1681 gttattcgtgtgtggtggttgcgttaaaacggtcaaatatgtaggaatttaactccggtlaca 1740
    |||
DB 1681 gttattcgtgtgtggtggttgcgttaaaacggtcaaatatgtaggaatttaactccggtlaca 1740
    |||

QY 1741 tcgcaattgaaacactcaggttactctgtccttgcgggtgtaggtggaagtcggtcgaag 1800
    |||
DB 1741 tcgcaattgaaacactcaggttactctgtccttgcgggtgtaggtggaagtcggtcgaag 1800
    |||

QY 1801 gtaacgacttcatcatatgataatgaaagacttaacgtaattagataaagtgtctctgttt 1860
    |||
DB 1801 gtaacgacttcatcatatgataatgaaagacttaacgtaattagataaagtgtctctgttt 1860
    |||

QY 1861 atgttcaaatgaaagcagccacaggttaacggtttagcgtgtgtgcatataacaggttaacta 1920
    |||
DB 1861 atgttcaaatgaaagcagccacaggttaacggtttagcgtgtgtgcatataacaggttaacta 1920
    |||
```

OY 1921 tggcgagaaatctcgatgaaggtcgatgtctattcttcgttgataa 1972
 |||
 DB 1921 tggcgagaaatctcgatgaaggtcgatgtctattcttcgttgataa 1972

RESULT 2
 AAA81456/C
 ID AAA81456 standard; DNA; 46593 BP.

XX AAA81456;

DT 04-DEC-2000 (first entry)

XX N. meningitidis partial DNA sequence gnm_4 SEQ ID NO:4.

XX Neisseria meningitidis; Neisseria gonorrhoeae; genome; immunogenic;
 antigen; vaccine; diagnosis; infection; antibacterial; identification;
 KM Meningococcus B; MenB; ds.
 XX Neisseria meningitidis.

OS Neisseria meningitidis.

PN WO200022430-A2.

PD 20-APR-2000.

XX 08-OCT-1999; 99WO-US23573.

PR 09-OCT-1998; 98US-0103794.

PR 30-APR-1999; 99US-0132068.

XX (CHIR) CHIRON CORP.

PI Frazer CM, Hickey E, Peterson J, Tettelin H, Venter JC;
 PI Maignani V, Galeotti C, Mora M, Ratti G, Scarselli M, Scarlato V;
 PI Rappuoli R, Pizza M;

DR WPI: 2000-318079/27.

XX Isolated nucleotide sequences of Neisseria meningitidis which can be
 PT used in the diagnosis and treatment of N. meningitidis infection and
 PT other Neisserial infections, for example, N.gonorrhoea -
 PS Claim 7; Page 260-274; 1760pp; English.

XX The present invention describes methods of obtaining immunogenic
 CC proteins from Neisseria genomic sequences. AAA81453 to AAA82414
 CC represent specifically claimed Neisseria meningitidis genomic DNA
 CC sequences; AAA81260 to AAA81303 and AAB25620 to AAB25663 represent
 CC Neisseria DNA sequences and their corresponding proteins; AAA81254 to
 CC AAA81259 and AAA81304 to AAA81321 represent PCR primers used in the
 CC isolation of Neisseria meningitidis DNA sequences; and AAA81322 to
 CC AAA81452 represent Neisseria meningitidis MenB polynucleotide ORF
 CC sequences, which are all used in the exemplification of the present
 CC invention. The nucleic acid sequences, protein sequences, and antibodies
 CC against them, can be used in the manufacture of a composition. The
 CC composition can be used as a medicament (or in the manufacture of a
 CC medicament) for treating, preventing or diagnosing infection due to
 CC Neisserial bacteria. For example, some of the identified proteins could
 CC be components of vaccines against Meningococcus B; against all serotypes;
 CC and/or against all pathogenic Neisseriae. Identification of sequences
 CC from the bacterium will also facilitate production of biological probes,
 CC particularly organism-specific probes. Attempts to make efficacious
 CC Meningococcus B vaccines have failed mainly due to antigen tolerance.
 CC Multivalent vaccines have also been tried but none have successfully
 CC overcome antigenic variability. The provision of further, complete
 CC sequences may provide an opportunity to identify secreted or surface
 CC exposed proteins that may be presumed targets for the immune system and
 CC which are not antigenically variable or at least more conserved than
 CC other more variable regions.

Sequence 46593 BP; 11355 A; 13195 C; 11355 G; 10687 T; 1 other;

Query Match 30.6%; Score 603.6; DB 21; Length 46593;
 Best Local Similarity 58.7%; Pred. No. 3.4e-139;
 Matches 1154; Conservative 0; Mismatches 779; Indels 33; Gaps 5;

OY 11 ttgctcaagtcgctcgatctctatgatgcacagcgaagcaattacatggtcaa 70
 |||
 DB 41623 TTCTGCCATTTGCAATCCGATTTGGATGAGCTAGCGCGCAAAACATTTGAGCATGGCGAA 41564
 OY 71 aaagtaactgaattatgaacaaacaaactactctcgatctgtagacaaagca 130
 |||
 DB 41563 GTGCTAACCCGAACCTGATGAAGAAACGAAACAGTTACAGTACCTGTAATCGGCTGAATAGCT 41504
 OY 131 ttagtgatttgcgtagaagttggttacttagaagaagtgacttagatgtagt 190
 |||
 DB 41503 TTGACCCCTTTGGGCAATCAACAACGGTTGTTCTGATGTTCCGGTTGCCAAAGCTCTG 41444
 OY 191 taattgaatcagcaactttagatagtagtaacataactatgctatttagtcgtag 250
 |||
 DB 41443 GCTTTGAATCTGAATTTTGTGACCTTTGTCGCTACCCACATCCGGAATTTTGGAGGCC 41384
 OY 251 ttaacccaactctggaataacataatgaatgaatgaagagatcaataagagcatttgat 310
 |||
 DB 41383 GTCAATGCTTCAGTGCATATGTCGACGAGAGAGAGAAACGCTTGAACGACCATGAA 41324
 OY 311 agcttcaagcaaacagctgctggtgaagtaacacttaatggagagca----aaatg 366
 |||
 DB 41323 TCCCTCAAAATCTTCTTACCCCTATCGCATTAAGCTTGAAGTGAAGAGTCTGAAATG 41264
 OY 367 gcaagtgctaaagagaataaagcaaacacgcgagtgtaaaagacacaaataact 426
 |||
 DB 41263 GCAGTGAAGAAAGAGATTTTCAACCAAAATCCGACGTTTCAGAAATCCCAAAAGATCACT 41204
 OY 427 aaagcgaatggaatggtgtgtcgtcgcgaatgcytaaaacgaagacgaatgctct 486
 |||
 DB 41203 AAAGCATGCATGATGTTGTCATCACTTAAATGCGGAACACTCAGAGAGAGTGGCTTTG 41144
 OY 487 tcaagccctattcgaagaacataagtaacgtgattaggaacaggtttccaagaacagat 546
 |||
 DB 41143 GCGCGTCCGTATGCCGAAAGAGTGTATGATGATGAGCCATCTTGGCAAAACCAATACC 41084
 OY 547 ggttacaagcatcattttt---agtgaatcggaatgaataaaagtggtgcatgtgt 603
 |||
 DB 41083 GATCATGATATTCCTTACTGTGAATCTCATCGGAAATCAGACGTGTGTTTATTTTG 41024
 OY 604 gttccacagatcgtgtgttctgtgtggttaaacgtgaactgttttaaacatgata 663
 |||
 DB 41023 ATTACGCTCATGAATGAGGTTTGTGTGCGCTTGAACGCCACGATGCTGAATAAGTTTGG 40964
 OY 664 aatgaatgaagaatgaagaagaagaatgtttccgttcaattgaattagttcgttct 723
 |||
 DB 40963 GCACAAAGTTCAAGAGATATCGGAATCAAGGATTTGAAGAGCAAAATCGTATGTTGGCAGT 40904
 OY 724 aaatcataacttttccaactcttgggaatgaataaatttcaacaaatcaagttat 783
 |||
 DB 40903 AAAGGTCTATGCGGTGTCAGACATGTTGTTGTAATGTGTTGCCAGTCCGTAATTTG 40844
 OY 784 gttgatactccctctgttgaagcaatgttgaatgttcaagtaattctatgtagtctat 843
 |||
 DB 40843 GCGGATACCCCAAAATGGAATGTGCTGCGACCTTTGACGAACCTCTTCCAAACGCTGT 40784
 OY 844 aaaaaaggggaatgtagtgtttagtatttaacaaattatcaacgaatgctg 903
 |||
 DB 40783 GAGAAACATGAATATGACGAATTCATGCTGATTTCCGGTTTGTCAATACCATGCGCT 40724
 OY 904 caaaagcagatttgaanaattatccattacacagaatagatgaatgaatgagc 963
 |||
 DB 40723 CAAGAAGCAGAAATGAATATTTGCTGCTATTCGTTGAGAAAGTGTGCGATTACGCT 40664
 OY 964 gaaagaaacgaagt---ttggattatattacgaactgtagcgaagatattatagat 1020
 |||
 DB 40663 CCCAATACACCGTTACGCTGAGGATACCGCTACGAACCCACATGCACTTGCAGTGTGGAA 40604
 OY 1021 aattatgtgtcgtcatttagaatctcaaggtttatcaagcagcagttgaaacactgct 1080

OY	121	tcatttgatcagcactctttagagtaagtaagtcacacatacctaibctgattttagtcgtga	250
Db	234563	GCCTTTTGAAATCTGGAATTTTGTGAGCGTTTGTCCGTACCCACATCCGGAAAGTTTGGAAAGCC	234504
OY	251	ttaacccaatctggtcaatlaacaatgtaagtaataaagcattaaagcatttgat	310
Db	234503	GTCAATGCTTACAGTGCANHTCCGACGAGAGCGAAGAAAACGTTGAGACAGCCATGAAA	234444
OY	311	agctcaaaagcaaacagtcgctggtgtaagtaacattaaatgtagagaca---aaatg	366
Db	234443	TCCTTCAATCTCTTACAGCCTTACGAGCAATTAAGCGTTGAAGTGAAGATGCTGAAATG	234384
OY	367	gcaggtgtcaagagataagaaccaaaatcgaggtgtataaagtlacacaaaaatct	426
Db	234383	GCAATAGGAAAAAGAAATTCACCAAAATCCGAGTGTTCAGATACCACAAAAGATCACT	234324
OY	427	aaagcgaatggaatgagttgtgcctcgcaaatggttaaaacgcaagaagcatgctctc	486
Db	234323	AAAGCGATGCGATGGTGGTCAACTCTTAANAATGGGAAGCTCAGGAACGATGCGTTTG	234264
OY	487	tcacgaccttattagaacaacatcgtaaacgtgtatgacacglttccaagaacagat	546
Db	234263	GGGGTCCGTATGCGCGAAAAGTCCGATATGGTATGAGCATCTTGCGCAACCAATACC	234204
OY	547	ggttacaagacatccactt---agtgatcgagaagtaaaaaagtgggcatgatt	603
Db	234203	GATCATGATATCCGTTACGTACGTGGAATCTCATCGGGAATTCAGAGTGTCGGSTTTTATTTTG	234144
OY	604	gtgcacacagtcgtggtcttlttggtggtgttaaacgttgactgtttaaactgtatta	663
Db	234143	ATTACGTCGTGAATAGGGTTGTGTGGGGTTTGAAACCCCAACGTCGAAAAGTTTGG	234084
OY	664	aatgaatgaagaagatggaagaaaagaatglttcgttcaatgagtttaactcggttc	723
Db	234083	GCACAGTGTCAAGATTCGGAATCAACAGATTTAAAGAGAAATCGATCTTTGGCAGT	234024
OY	724	aaatcblcaactcttttccaaactcttgggaattaaatlttaacccaagaatlcagtlat	783
Db	234023	AAAGTGTGATGGCGTGTCAAGCATTTGTGTAATGTGGTCCCAATGCGCGTAATTTG	233964
OY	784	ggtgatctccctctgttgagcagttaatggttcaatcaatctatagatgtgcttat	843
Db	233963	GGCGATACCCCAAAATGGAATCTCTCGGACCTTTGACAGAACTTTCACACGGTAT	233904
OY	844	aaaaaggggagtagatgtgttactttagttataaacaattatlaacagatctcg	903
Db	233903	GAGAAACATGAAATTTGACAGAAATCCATCTGTGTATTCGGGTTTGTTCATATACCATCGT	233844
OY	904	caaaagccagtatggaanaaattaatccatlacacagaattagataatgaatgaatgagc	963
Db	233843	CAAAACCGAGAAATTTGGAAGTATTCTCTCATCGTGTGAGAACGTAATGGCGATTGAGCT	233784
OY	964	gaagaagaaacagt---ttggatataatttagaacctgatacggaagtattattagat	1020
Db	233783	CCCAATATCACGGTTCAGCTGGGAATACCCTGACGAACCGATCGACTTGCAGTGTGGAA	233724
OY	1021	aattatgtgtcttatttagaattccagtttatcaagcagcagttgaaacctgtcg	1080
Db	233723	TATCTGTTGCCCTATTTAAGATCTGTGTTTATATAGCGCTTGAAGCCCAATATGGCA	233664
OY	1081	tcgtagcaagccgctctgaatggttcgcacataaagcagcaacagataacgagtlaactta	1140
Db	233663	TCCGAACGACGACGCCGATGTTGCCATGTAACCTCAACAGACATCAAGGCAATGGCC	233604
OY	1141	attaatggttaacgttagctctataaacaagaactcgttcaagcaagttataaataatla	1200
Db	233603	ATCAAAAGGTTCGCTTTGGTATTAACAAATCGGCTCAACCTGCGATTACACGGAATTG	233544
OY	1201	aatgaatgtgttcggtgtgcagcagcaattaa-----caatag	1240
Db	233543	TCAGAAATTTGACAGAGTCCGGCGCGCTGTCTGATGACCCGCTGTGAATCTCTTACGGAATTA	233484
OY	1241	agagtcggtlaatgccaactggaanaaattgttacaacatcagtggtgttattgagcttga	1300

Db 233483	GGATACGATTAATGAGCCCAAGGCCAAATTCGTACAAATTAATGCGTGGCTTTTAACTGGGA	233424
Qy 1301	attccacaagaatcagttaccacaagaatataatgaccttaa---tggtgaacaggttt	1357
Db 233423	ATTTCACAGCAGACATGATTTCCGGCCGTTTACGACGCTTTGAAATTAAGCAAAACGGTCT	233364
Qy 1358	agttactgaagttaacaacaacaataagttggtggtgtagttccgtgatacgcaatggatc	1417
Db 233363	GACTTTGGAGATCCCAACGACTTTTGGCGCAGCGGGGTGTCGTGCCATTCGATGGGCGAC	233304
Qy 1418	atctgattgattaaacggggtttaagcgtaaacaaatagagaataaaccaaattctgttcc	1477
Db 233303	CTCGACGCTTTTGAACCGCGGCATAGCTGTGAGCAATACCGGTGCGCCATTACTGTGCC	233244
Qy 1478	agtgggaacggaanaacatgggtcgtatcatgaagatacttggtgaaccaatcgatagca	1537
Db 233243	GGTAGCTAAAGCTACGTTTGGAGCCCATTTTCGATGATTTGGGAACGCTTTGACGAAGC	233184
Qy 1538	aggltgaatcgtgtgcagaagaatggtgtcatctatccacgltgcgcaccaagtaaga	1597
Db 233183	AGGTCCGATTCGATACCGCAGACAGCCGTCGCATCCACCAAGCTGCCCTTAAGTTGACGA	233124
Qy 1558	acaatcaacagttactgaactttagaagaagggaataaagtatcgacttagttgtcc	1657
Db 233123	ACTGCTTCCCAACACCGAATTGCTCGAAGACGGGCATTAAAGTGAATTGACTTGCTGTGTC	233064
Qy 1658	gtttcggaagaagggtgaagaagttatttcggtgtgtgcgggtgtcgtaaacgttcaa	1717
Db 233063	GTTTCCCAAGGCGGTTAANAGTAGTCTGTTTGGCGGTGCGGTGTGGGTAAACCGTGAA	233004
Qy 1718	tatgattgaattaaatccgtaacatcgcgaatagtagcactcagttactgttcttgcgg	1777
Db 233003	CATGATGGAATTGATTCACAAACATTCGCGCAAAAGCGCACAGCGGCTTGGCCGTTTCGAGG	232944
Qy 1778	ggttagtgtagcggtacggtgaaggttaacgactctcatatgagatgaagaacttaacgt	1837
Db 232943	CGTGGGTAGCGGTACCCCGCAAGGTAAAGACTTCTACACAGAGATGAATTTCCAAACGT	232884
Qy 1838	attgattaaagtgtctctgttctttagtgcgaatgaacgagccaccagtagtaacgtttacg	1897
Db 232883	ATTGGATTAAGTGGCCATGGTTTACGGTTCAGATTAAGAACTTCGCGGCAACCGTTTGGC	232824
Qy 1898	tgtgcatctaaccaggtctaactatagcgcgaaaaatctcgtgatgaa	1943
Db 232823	CGTCCGATTTGACCGGTTTGACCAATGGCGGAATACTTCCGTGACGAA	232778
RESULT 4		
AAA81489/c		
ID	AAA81489 standard; DNA; 837096 BP.	
XX	AAA81489;	
AC		
XX		
DT	04-DEC-2000 (first entry)	
DE	N. meningitidis partial DNA sequence gnm_37 SEQ ID NO:37.	
XX		
XX		
XX	Neisseria meningitidis; Neisseria gonorrhoeae; genome; immunogenic;	
KM	antigen; vaccine; diagnosis; infection; antibacterial; identification;	
KW	Meningococcus B; MenB; ds.	
XX		
OS	Neisseria meningitidis.	
XX		
PN	WO200022430-A2.	
PD		
XX	20-APR-2000.	
XX		
PF	08-OCT-1999; 99WO-US23573.	
PR	09-OCT-1998; 98US-0103794.	
PR	30-APR-1999; 99US-0132068.	
XX		

0Y	847	aaggggaagtaagatgctgtgtgtgtattagtttaatacaaatattataacaacgatgtcgaa	906
Db	1225	gaagaaagaattatgataatataataatactatagtcattatgtcaaggttcttgaaac	1284
0Y	907	aagcagtaattggaataatcatccatccaggaattagataatgatgtaatgaagcgaa	966
Db	1285	aagcctactctagacaagatattacattcttcacagagatcttcagtaaggggatggt	1344
0Y	967	agaaacaagttcggattatattcgaacctgatgtcgaaagttatttgaataatta	1026
Db	1345	-----catgttcttcttatagttaatttggccagataaagaactctatcttaagttaac	1398
0Y	1027	ttgtgtcttatttagatctcaggttttcaacgscggttgaanaacctgtcttcag	1086
Db	1399	ttgcctcaatatgttggagtttgaattcttcgcgaacaattatagaacgaanaagtgag	1458
0Y	1087	caagcgcctcgaatbgtcgcgcataagaaagacaacagataacgcgagtaacttaatt	1146
Db	1459	catgcacacagctatgactgagtgaaanaatgccactgatattgcaactcgaacttatg	1518
0Y	1147	gagttacaagttagtctatacaagcttcgtcaagcagatattacaatggaattaaatga	1206
Db	1519	gacttaccattagataataacaagcgagaaacagacgaanaattaccacacaataattcaa	1578
0Y	1207	atctgtccggtgagcaagaactta-----caatlagaggtatcgtaattgccaact	1259
Db	1579	atgtgtgtgtgtccgagagcgttcgtataataatttaagaagagaaatatagcatggaatt	1638
0Y	1260	ggaanaaatctgacaacalcacgtgtcggttatgtacgttgaattcccaagaatgcagta	1319
Db	1639	ggccggtttaactcacaagattatggtgtccctgaattatgatttcgatttgaacatacgaagt	1698
0Y	1320	ccaagaattatgatagtccttaaatgttga-----aacaggttta	1358
Db	1699	ccttaaatataacgcgccttgattatgtatbgtccaaagaagaaggtacaaatacaacta	1758
0Y	1359	gtacttgaagtccaacaacaatlagtbtgtgtgtatgttcgtctgatccgaatbgtgata	1418
Db	1759	acattagaagtgtgcgtgcacattagtgtagcagcagtggttcgtacaaattgcagtattcca	1818
0Y	1419	tcgtatgattaaacgcggttttaagcgtaacaacatagataaaccattctgttcca	1478
Db	1819	actgatgtgtgtccaaagaagcagatgagctgtataaagaataacggcnaaagaattatgttacct	1878
0Y	1479	gtgggaacgcaaaacatbgtgtcgtacatacgtaaacgttatbgtgtgaacccaatcgatgcaa	1538
Db	1879	gtgtgtgtagcaaaacatlagtctgttatttaagtgtactagtggtgaacaacattgaaccttaa	1938
0Y	1539	gtgtgaattcgtgtcgaagaaggaattg---gtctattcacgcgtgtgcaccacaagttaatga	1595
Db	1939	gaagaataattgtgattctgttcgcgcgcgatacctcatccatcgtcacaacacgaacttgat	1998
0Y	1596	gaacaactacacgtaactgtgaacttttagaacaacgggataaagttatgcacttattggt	1655
Db	1999	gaacttccaacagaagtcaaatcttttagaacaacggtataaagtatgattactactaga	2058
0Y	1656	ccggttcgcgaaggggttaagtagttatttcgltgtgtcggtgtcgcgttaaacgcgtc	1715
Db	2059	ccttatcttaaaagtgtgttaaaatcggatattgttcgtgtgtgcgcggtgtgtatgaacaagta	2118
0Y	1716	aattgtatgaattaaatccgttaacatccgcgaatgtgaacatcgaactcgttactcgttcg	1775
Db	2119	ttaatccaagaattaatataacaacatccgtccaagaacacggtgtgattcttcgtatccgc	2178
0Y	1776	ggggtaaggtagacgtacgcgtgaaggtaacgacttcatacgtatgagatgaaagacttaac	1835
Db	2179	gggttaggtgaacgtactcgttgaaggtaacgatttatcttactcgaatatgtgtacagtggt	2238
0Y	1836	gtattagataaagtgctctctgtttatgtgtcaaatgaagaagcaccacagtgtaacggttta	1895
Db	2239	gtaatctagaanaacagccatgattatctcgggtcaaatgaaatgagccacactgtgtgtacagttg	2298

Oy 1886 cgtttgcatlaaaagccttaactatggcggaanaattccfgrtgaagtgcgtgatgc 1955
| | | | | | | | | | | | | | | | | | | | |
Db 2239 cggttcgattacatttggtcaacaacgycubgaataatlcccgfgaacgaacaagaftcaaac 2350
| | | | | | | | | | | | | | | | | | | | |
Oy 1956 ttattctt 1963
| | | | | | | | | | | | | | | | | | | | |
Db 2359 gtattatt 2366

RESULT 7

ID AAX13153 standard; DNA; 7750 BP.

AC AAX13153;

DT 19-MAR-1999 (first entry)

DE Enterococcus faecalis genome contig SEQ ID NO:216.

KW Enterococcus faecalis; contig; detection; Enterococcal infection;

XX

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120	121	122	123	124	125	126	127	128	129	130	131	132	133	134	135	136	137	138	139	140	141	142	143	144	145	146	147	148	149	150	151	152	153	154	155	156	157	158	159	160	161	162	163	164	165	166	167	168	169	170	171	172	173	174	175	176	177	178	179	180	181	182	183	184	185	186	187	188	189	190	191	192	193	194	195	196	197	198	199	200	201	202	203	204	205	206	207	208	209	210	211	212	213	214	215	216	217	218	219	220	221	222	223	224	225	226	227	228	229	230	231	232	233	234	235	236	237	238	239	240	241	242	243	244	245	246	247	248	249	250	251	252	253	254	255	256	257	258	259	260	261	262	263	264	265	266	267	268	269	270	271	272	273	274	275	276	277	278	279	280	281	282	283	284	285	286	287	288	289	290	291	292	293	294	295	296	297	298	299	300	301	302	303	304	305	306	307	308	309	310	311	312	313	314	315	316	317	318	319	320	321	322	323	324	325	326	327	328	329	330	331	332	333	334	335	336	337	338	339	340	341	342	343	344	345	346	347	348	349	350	351	352	353	354	355	356	357	358	359	360	361	362	363	364	365	366	367	368	369	370	371	372	373	374	375	376	377	378	379	380	381	382	383	384	385	386	387	388	389	390	391	392	393	394	395	396	397	398	399	400	401	402	403	404	405	406	407	408	409	410	411	412	413	414	415	416	417	418	419	420	421	422	423	424	425	426	427	428	429	430	431	432	433	434	435	436	437	438	439	440	441	442	443	444	445	446	447	448	449	450	451	452	453	454	455	456	457	458	459	460	461	462	463	464	465	466	467	468	469	470	471	472	473	474	475	476	477	478	479	480	481	482	483	484	485	486	487	488	489	490	491	492	493	494	495	496	497	498	499	500	501	502	503	504	505	506	507	508	509	510	511	512	513	514	515	516	517	518	519	520	521	522	523	524
--	---	---	---	---	---	---	---	---	---	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

XX

XX

XX

PR 06-MAY-1997; 97US-0044031

XX

XX XX

[illegible]

XX

PT - used to develop products for the detection of Enterococcus and for

PT infection.

PS Claim 1; Page 1098-1102; 2084pp; English.

CC A computer readable medium has been developed which has recorded on it

CC AAX12938 to AAX13919 represent these nucleotide sequences which are

CC system can identify fragments of the *Enterococcus faecalis* genome with

CC of *Enterococcus faecalis* in samples. They can also be used for

CC progression of disease, and for identifying agents which can be used to modulate the severity of Enterococcus faecalis or

CC another related organism, in vivo or in vitro. In particular the

CC can be used in vaccines to prevent or attenuate an Enterococcal infection

XX
Cecilia 7760 PD. 3530 A. 1373 C. 1531 C. 3315 E. 10 October

Query Match	17.08;	Score 335.2;	DB 20;	Length 7750;
Post Local Similarity	55.38;	Prod NO 3	80-73;	

Matches 836; Conservative 0; Mismatches 584; Indels 91; Gaps 6;

QY 547 ggttacaagcatccatttttagtggatcgcgaagtaaaaaagtgggcattgtgtg 606

Db 2082 gggaactatcatgtgatgttagcacaacgacctgtcaaaaaaaaaacaggtatatcgttattc 2141

FT	CDS		complement (39873..40514)
FT		/tag=	x
FT		/label=	MG034
FT		/note=	"Previously identified as MORF-20101, the encoded protein shows 48.13 percentage identity to thymidylate kinase (tdk) from B. subtilis"
FT	CDS		40543..41787
FT		/tag=	l
FT		/label=	MG035
FT		/note=	"Previously identified as MORF-20102, the encoded protein shows 30.71 percentage identity to histidyl-tRNA synthetase (hss) from Mycobacterium leprae"
FT	CDS		complement (44751..46277)
FT		/tag=	m
FT		/label=	MG038
FT		/note=	"Previously identified as MORF-20105, the encoded protein shows 46.83 percentage identity to glycerol kinase (gpk) from E. coli"
FT	CDS		complement (46268..47422)
FT		/tag=	n
FT		/label=	MG039
FT		/note=	"Previously identified as MORF-19831 and MORF-20106, the encoded protein shows 43.20 percentage identity to glycerol-3-phosphate dehydrogenase (GUT2) from S. cerevisiae"
FT	CDS		49377..49643
FT		/tag=	o
FT		/label=	MG041
FT		/note=	"The encoded protein shows 48.86 percentage identity to phosphohistidinoprotein-hexose phosphotransferase (ptsH) from Mycoplasma capricolum"
FT	CDS		50060..51520
FT		/tag=	p
FT		/label=	MG042
FT		/note=	"Previously identified as MORF-19832 and MORF-20108, the encoded protein shows 41.92 percentage identity to spermidine/putrescine transport ATP-binding protein (ptaA) from E. coli"
FT	CDS		51525..52382
FT		/tag=	q
FT		/label=	MG043
FT		/note=	"Previously identified as MORF-20110, the encoded protein shows 26.51 percentage identity to spermidine/putrescine transport system permease protein (ptcB) from E. coli"
FT	CDS		52366..53220
FT		/tag=	r
FT		/label=	MG044
FT		/note=	"Previously identified as MORF-20111, the encoded protein shows 29.45 percentage identity to spermidine/putrescine transport system permease protein C (ptcC) from E. coli"
FT	CDS		54658..55605
FT		/tag=	s
FT		/label=	MG046
FT		/note=	"Previously identified as MORF-20112, the encoded protein shows 36.60 percentage identity to sialoglycoprotease (gcp) from Pasteurella haemolytica"
FT	CDS		complement (56970..58310)
FT		/tag=	t
FT		/label=	MG048
FT		/note=	"Previously identified as MORF-19834, MORF-20114 and MORF-20115, the encoded protein shows 43.02 percentage identity to signal recognition particle protein (fth) from B. subtilis"
FT	CDS		58117..59079

Protein	Accession	Gene	Organism	Protein	Accession	Gene	Organism
FT	1	/*tag=	u	FT	1	/*tag=	u
FT	1	/label=	MG049	FT	1	/label=	MG049
FT	1	/note=	"Previously identified as MORF-20114 and MORF-20115, the encoded protein shows 44.78 percentage identity to purine-nucleoside phosphorylase (deob) from E. coli"	FT	1	/note=	"Previously identified as MORF-20114 and MORF-20115, the encoded protein shows 44.78 percentage identity to purine-nucleoside phosphorylase (deob) from E. coli"
FT	1	59083..59754		FT	1	59083..59754	
FT	1	/*tag=	v	FT	1	/*tag=	v
FT	1	/label=	MG050	FT	1	/label=	MG050
FT	1	/note=	"Previously identified as MORF-20117, the encoded protein shows 83.03 percentage identity to deoxyribose-phosphate aldolase (deoc) from Mycoplasma pneumoniae"	FT	1	/note=	"Previously identified as MORF-20117, the encoded protein shows 83.03 percentage identity to deoxyribose-phosphate aldolase (deoc) from Mycoplasma pneumoniae"
FT	1	complement (64698..65731)		FT	1	complement (64698..65731)	
FT	1	/*tag=	w	FT	1	/*tag=	w
FT	1	/label=	MG056	FT	1	/label=	MG056
FT	1	/note=	"Previously identified as MORF-20122, the encoded protein shows 30.25 percent identity to the protein disclosed in GB:D6185.99 from B. subtilis"	FT	1	/note=	"Previously identified as MORF-20122, the encoded protein shows 30.25 percent identity to the protein disclosed in GB:D6185.99 from B. subtilis"
FT	1	complement (65713..66249)		FT	1	complement (65713..66249)	
FT	1	/*tag=	x	FT	1	/*tag=	x
FT	1	/label=	MG057	FT	1	/label=	MG057
FT	1	/note=	"Previously identified as MORF-20123, the encoded protein shows 38.90 percentage identity to the protein disclosed in GB:D26185.104 from B. subtilis"	FT	1	/note=	"Previously identified as MORF-20123, the encoded protein shows 38.90 percentage identity to the protein disclosed in GB:D26185.104 from B. subtilis"
FT	1	81047..82597		FT	1	81047..82597	
FT	1	/*tag=	y	FT	1	/*tag=	y
FT	1	/label=	MG067	FT	1	/label=	MG067
FT	1	/note=	"Previously identified as MORF-19845, the encoded protein shows 28.84 percentage identity to glutamic acid specific protease (Spase) from Staphylococcus aureus"	FT	1	/note=	"Previously identified as MORF-19845, the encoded protein shows 28.84 percentage identity to glutamic acid specific protease (Spase) from Staphylococcus aureus"
FT	1	91065..91919		FT	1	91065..91919	
FT	1	/*tag=	z	FT	1	/*tag=	z
FT	1	/label=	MG070	FT	1	/label=	MG070
FT	1	/note=	"Previously identified as MORF-20136, the encoded protein shows 34.8 percentage identity to ribosomal protein S2 (rps2) from Spirulina plantensis"	FT	1	/note=	"Previously identified as MORF-20136, the encoded protein shows 34.8 percentage identity to ribosomal protein S2 (rps2) from Spirulina plantensis"

Query Match	15.1%;	Score 297;	DB 18;	Length 580073;
Best Local Similarity	49.68;	Pred. No. 4.5e-63;		
Matches 979;	Conservative	0;	Mismatches 935;	Indels 61;
				Gaps 6;

[illegible]

OY	409	agtcacaaaaaattacttaagaagatgaaatggttgcgtcccggaatatgctaaacg	468
Db	505054	TCCACCACTTAAGATAACTAATGCAATGAATAAATGGTGTGCACGCCCTAAGTTATTAAAGTTTC	504995
OY	469	caagaacgcagtctccttccacgcccttatccgaacaatacgaacgtgatagccac	528
Db	504994	AAAAAACAGTTTTCAAGAAATTAGTTTGTTTTTAATGAATTTATFAAGCGTTGTGGCAA	504935
OY	529	gttcccagaacacgatgtgttacagcatccttlttgtagatcgcggaataaaaaa	588
Db	504934	GT-----AGTTGTTTTCTTTAAAACAACCAAAAAAGAACCAGTAACCAAAAA	504887
OY	589	gtggacatgatgtgtgtccacagatcgcgtgtcttgtgtgctttaaacgtgaacctg	648
Db	504886	ACTTATGGAATAATGATGAGTGTCTGTTTGAAGCTTTGTGGACAGCATATATGGAACATG	504827
OY	649	tctaaccctgtatctaataatgaaatgaaagaatgaaagaagaagaatgttccgttcaatg	708
Db	504826	AATVAGTTAATTAAGAAAGCTAATTTTAAAGCTGATGATGAATAATCTTTTTTTAGGTAGAGAAA	504767
OY	709	agtttaacgtgttccaatctcatacatttttccaatcttggaaattaacatttcaacc	768
Db	504766	AACCAATCATTTCGAATTAATAATAGCAATATATACCTCTGTGGATTATTTGATATC	504707
OY	769	caagatccaggatgtgtgtaactccctcgtgttgacagttaatggttcagtaacct	828
Db	504706	CAGATCGTGATATTT-----AATTTGATTAATGTCAAAAGATATTTGATCAG	504659
OY	829	atgtatgatctgataaaaaaggggaagtgaatgattgtgattagttataacaattt	888
Db	504658	ATTATGATGTGATTTAAAGAGTTTAAACTTGATGCAATTTGTATGGTTTACCTTAATATTT	504599
OY	889	attacacgcagtcgcgcaaaagccagtaatbtgaaanaaaltaatlccatcacagaattagat	948
Db	504598	AAAAACATTAATCAACAACATCTCACGCTCTTCAAGTTTTCGATGTTTGGAAAGT	504539
OY	949	aatgataatattaggcgaagaagaaacaaagtttggattatataattcgaacctgtgcga	1008
Db	504538	TTTTAAACCTTTAAATCCGGTGTGTAACGATCAACAACTTTATTTTACGCCAGATVCAAGCC	504479
OY	1009	gtattctatgataattatcttgcgttcgtatctagaatcctcaggttatacaagaacagtt	1068
Db	504478	ACGATATTATTATTAATTACTCCACACTTTTGTGATGTGGCTCTGTATGTGTGGCCTTTGTT	504419
OY	1069	gaaaacctgccttcctgcgcaagccgcgtcgaaaygttcgcataagaaagcagcaacagataac	1128
Db	504418	GAAACTAGTTATGTGANTGANTCAGCTTTCAGCAAAATGCAATGGAACCTGCTCAAAGAAT	504359
OY	1129	gcaggtacttaattaaatgtgttaacggttgcgtctataacaagaactcgttaagaagtatt	1188
Db	504358	GCTAAGATTACTTGATTAATACACTTTTCAANTTTAAACAAGCTAAGACAAAACCTGTATT	504299
OY	1189	acaaatgaatlaaatgaatctgttcgcpgtgcagcagcaatttaacaacatagagatcg	1248
Db	504298	ACGAACAAGATTATTGAAAGTTATAGAGAGATGATGAATTAATTATGATTAATAAAACAAACCT	504239
OY	1249	taatgycacactgaaanaattgtacaacaatcoggtgcggtalttgacgyltgaatlcceao	1308
Db	504238	AA----CATATGTTAAAGTTCCACCAATCTCATTTGTGCTGTAGTTGAATGTTATCTTTACAG	504183
OY	1309	aa---ggtgcagatcccaaagtataigaicccctaagtgtg-----aaacaggtt	1356
Db	504182	AAAGTTAAACATTACCTAGAGTTTATATATGTTTGAAGTTACAACTRAAAAAAAGTGAGC	504123
OY	1357	tagtacttgaagttcaacaacaatlagtvtgvtgvtgtagtgcgtlatcgcaatvgygat	1416
Db	504122	TTTTTTTGAAGCAACCCAATTAATAGGTATAGGTATGACATTGTTGCGTCAATTGCATVTAGGTC	504063
OY	1417	cattcgtatgattaaacgcggttttaagcgtlaacaatacgaataaccaatttccgttc	1476
Db	504062	CTACAGAGGATTTAGCACGCTAATGTTTAAATTTACTAATCAATTCACATVTAGAGGTAC	504003

OY	1477	caatgggaaacgaaaaccttggctgcgtacatgaacgattggctgaaacaaatgcgaagc	1536
Db	504002	CTGTTGGCAAAAACGTATTGGGGAAGATGTTCAATGTTTAGGTGAACCATTCATGATGAA	503943
OY	1537	aaggatgaatcaggtgcgaagaagatgctatctcaacgctgycgaaccaagttatgaag	1596
Db	503942	AAGAACCATTGCCAAAAAACCCAAAGCTATCATTCATCCATCTAAACCAACCGCTTTGGATG	5038833
OY	1557	aacaactcaacagctacggaactcttttagaaaacgggaattaagttatccgacttgcttgc	1656
Db	503882	AACAACCAAAATACGTTGATAATTTTGAACAGCAATAAAAAGTAATGATCTTTTAACTC	503823
OY	1657	cgttcgcgaaggggtgaagaatgaattatcgtgtgctgcggctgcgtgtaaaacgctca	1716
Db	503822	CTTACGTTTAGGGGGGGGTAATTCGTTATTATTGGAGAGCGTGCGTTGGGAAAAACGTTT	503763
OY	1717	atatgatggaatatactccgtacaatcgcgaattgagcaactgaagttactctgtctttgcg	1776
Db	503762	TGGTGCAAGAAATTAATTCATTAATCCCAAGAACCTTCTGGTTTAAGTGATTTGGCTG	503703
OY	1777	gggtgaagtgagcgaacgcgtgaagtgaaagactctcactgaagatgaagaagactcaag	1836
Db	503702	GAGTTGGTGAAGAAACAAGACAGATTAATGATCTTTACTTGAATAATTCATCAAGGTGGG	503643
OY	1837	tattagataaagtgctctcttgtttatgtgtcaaatgaaacgagccacaagtaacggtttac	1896
Db	503642	TGATGTGATTAACACGTTTATGTTTGGCCAAATGATGAAACCAACGAGAGCTAGAAATGA	5035833
OY	1897	gtctggtcattcaacgcgttactactatgctgcggaataaattccgtgatgaagtcgcgta	1951
Db	503582	GAGTGTGCTTTTAACTGCTTTAAACAATGGCAAGATATATTTTCTGTGATCAAGATAAATCA	503528

RESULT 10
 AAC79635
 ID AAC79635 standard; DNA; 531 BP.
 XX
 AC AAC79635;
 XX
 DT 08-FEB-2001 (first entry)
 XX
 DE Virulence gene #44.
 XX
 KW Virulence gene; antibacterial; vaccine; bacterial infection;
 KW septicemia; bronchopneumonia; rhinitis; wound infection; ss.
 XX
 OS Actinobacillus pleuropneumoniae.
 XX
 PN W0200061724-A2.
 XX
 PD 19-OCT-2000.
 XX
 PF 06-APR-2000; 2000MO-US09218.
 XX
 PR 09-APR-1999; 99US-0128689.
 PR 10-SEP-1999; 99US-0153453.
 XX
 PA (PHAA) PHARMACIA & UPJOHN INC.
 XX
 PI Lowery DE, Fuller TE, Kennedy MJ;
 XX
 DR WPI; 2000-647422/62.
 XX
 PT Attenuated Pasteurellaceae bacteria comprising mutations in virulence
 PT genes, useful as a live attenuated vaccine against bacterial infections
 PT -
 XX
 PS Example 5; Pages 238-239; 322pp; English.
 XX
 CC The family Pasteurellaceae encompasses several pathogens that infect a
 CC wide variety of animals. The present invention relates to virulence genes
 CC from Pasteurellaceae. The present sequence is one such virulence gene.
 CC The present sequence may be mutated in order to produce an inactive gene.
 CC

QY 491 gccctattcagaacaacacgtaacgtat-----tagccagctttccaagcaaga 544
 DB 3608 aagtttcctcagaagaagtcgtaaacctttgcagatactcttccttcgttaacgtgagctg 3667
 QY 545 ttggttacgaagcctcatttttagtgatcgcaagtaaaaaaagcgagctatggttg 604
 DB 3668 gtggttcaactaacccagtggtgattagccgtctcgtgaagaagacagctatatacgttca 3727
 QY 605 tctccacagatcgtgtgcttcttggtgctgtaaacgltgaactgttttaaacgttataa 664
 DB 3728 tcaactcagaccggtgtgtgtgtgaggtatatactccctattttgaaagcgcttatg 3787
 QY 665 atgaaatgaagaatggaagaaagaaagatggttcogtccaattgagtttaactggttca 724
 DB 3788 agtctgaagaagaatcaccacccagacggttaaaagttttgaaatgactgtatcggtgga 3847
 QY 725 aactatacaacttttccaactcttggaatttaaaatttaaccagaattcagttatg 784
 DB 3848 tgggagctgatacttcttaaggtcgcggtatccaactttatgatacgtggtgtg 3907
 QY 785 gtgatacctcctctgttgacggttaattgtgtcagttcaattctatgattgattgata 844
 DB 3908 cagaccacactagcttggatccaagttcgttaagattatttcaaaaactgttgaatgtacc 3967
 QY 845 aaaaaggggaagatagatgtgtgtattgattatatacaaaattatatacagatgtgc 904
 DB 3968 aaaaatgaactcttgtagcttattgtttgttcatacaaccacatgtaacgttaacca 4027
 QY 905 aaaaagcagatttgaaaaatttaattccattaccagaattatgataatgaaattgagcg 964
 DB 4028 gtcaaatgtgtggaacaaatgcttcgactgtgtgactgtgataccaatga---agcgg 4084
 QY 965 aaagaacaagcttggagttatattacgaacctgatacgaaagtattatagataatt 1024
 DB 4085 atgaagagatcagcgttgaactttgaaatgaaacagccgagaagaatcttgagcagt 4144
 QY 1025 tatgtgtgtattatagatcgaagttatataacgacagttgaaacctgtctcg 1084
 DB 4145 tctgtcctcagtttgcaagaagatgatttaacgttgcattacagatgcgaagacagctg 4204
 QY 1085 agcaaacgcctcgaatgagtcgcacatgaagaacgacagacagataacgaggaattatata 1144
 DB 4205 agaatgtctgggacagacagccatgcaaacgacgacagataatgctaaagaatcatca 4264
 QY 1145 atgattacagttagttctataaacaagctcgttcaagaagattatacaaatgaaattaat 1204
 DB 4265 atgatttgcaattcagttataacacgtgcagacagcggtgattacacaagaattacag 4324
 QY 1205 aattgttgcggtgacagacgaatttaacaatatagagatc----- 1246
 DB 4325 aatcgttagcaggtgctagttccttagaataagctctagtcacgtctgatatgaatagaa 4384
 QY 1247 -----ggttaag 1253
 DB 4385 cttagagactagtgtgacgacgaacccgacgtatcttataatagaagaagaagagag 4444
 QY 1254 gcaactggaanaattgtacaacatcaggttgcgttatitgacgttgaattccacaagaat 1313
 DB 4445 agtcaagtaaaatgtcaggttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgc 4504
 QY 1314 gcaagtcacaaagtatataatgcttaaatgtg-----aaaca 1352
 DB 4505 gaaaacactcttcgtagattacaatgacattgctgtctacaaaaatgacgaagaanaaaa 4564
 QY 1353 ggtttagtactgaagttcaacaacattaggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1412
 DB 4565 aaaaatgcttctgaagtagcctgt 4624
 QY 1413 ggtatctgtagtgaataaagcggtttaaagcgttaacaatacgaataccaactttc 1472
 DB 4625 gatacaacagatgt 4684

QY 1473 gttccagttggaaacgaanaacattggtgtgtatcatgaacgttattgttgaaaccaatgat 1532
 DB 4685 gtaccagttggttaagaagaacttggagcgtgtcttcaacgcttttggagatatacatgtgac 4744
 QY 1533 gaaaggtgaanaatcgtgtcaga---agaagattgtgtctatccacgttcgcgaccacaaat 1589
 DB 4745 ttggaagctccttttaccagaagacgacagagcgtacgaaccaattatataaaaagttccaact 4804
 QY 1590 tatgaagaacaactcaacagttactgaacttttagaagaacggaaattaaagttatcagctta 1649
 DB 4805 ttgtatgattgtgtactccttcttgaatcccttgaaacagggaatcaagttattgacctt 4864
 QY 1650 gttgtccgtttgcgaagaagggttaagtagttattcgtgtgtgtgtgtgtgtgtgtgtgt 1709
 DB 4865 ctgtcccttaccattaaagtgtgtaagttggaacttctcgtgtgtgtgtgtgtgtgtgtgtgt 4924
 QY 1710 accgtcaataatgataatcgcgttaacatcgcgttaacatggaacacgacgttactctgtc 1769
 DB 4925 actgtcttaatccaagaatgtattatcaacatgacacatgcccagaagacggtgtgtattcagta 4984
 QY 1770 ttgcgggggtgaggtgagcgtacggtgaaagtaacgacttctatcatgagatgaagaac 1829
 DB 4985 ttgtcgtgtgtgtggaagcgtactcgtgaggtggaatgaccttactcgtggaatgaaaga 5044
 QY 1830 tctaaagttatagataagttgtctctgtttatgtgtcaatgaaagacgagccacaggttaac 1889
 DB 5045 tcaagggttactcgaanaaacacgacagcgttcttgttcagatgaatgagccacagagaa 5104
 QY 1890 cgtttagcgttgacatcaacgacttaactatgctgcggaanaaatccgtgtatgaagtcgt 1949
 DB 5105 cgtatcgtgtgtgtccttactgtgttgacaatcgtcgaataactcgtgtatgtgtgaagtc 5164
 QY 1950 gatgtctattctcgtc 1967
 DB 5165 caagacgtgtcttctctt 5182

RESULT: 12
 AAV29571
 ID AAV29571 standard; DNA; 4815 BP.
 XX
 AC AAV29571;
 XX
 DT 04-AUG-1998 (first entry)
 XX
 DE L. lactis soluble part (F1) exhibiting ATPase activity encoding DNA
 XX
 KW Biomass production; uncoupled ATPase; F0F1 ATPase; membrane bound;
 KM F1; Lactococcus lactis; ds.
 XX
 OS Lactococcus lactis.
 XX
 FH Location/Qualifiers
 FT 26..553
 FT CDS
 FT /tag= a
 FT /product= "ATPase subunit".
 FT /gene= "atpH"
 FT /note= "delta subunit of the F1 portion of F0F1 ATPase"
 CDS 742..2244
 FT /tag= b
 FT /product= "ATPase subunit"
 FT /gene= "atpA"
 FT /note= "alpha subunit of the F1 portion of F0F1 ATPase"
 CDS 2260..3129
 FT /tag= c
 FT /product= "ATPase subunit"
 FT /gene= "atpC"
 FT /note= "gamma subunit of the F1 portion of F0F1 ATPase"
 CDS 3301..4710
 FT /tag= d
 FT /product= "ATPase subunit"
 FT /gene= "atpD"
 FT /note= "beta subunit of the F1 portion of F0F1 ATPase"

XX WO9810089-A1.
 XX 12-MAR-1998.
 XX 08-SEP-1997; 97MO-DK00373.
 PF 06-SEP-1996; 96DK-0000963.
 XX (JENS/) JENSEN P. R.
 XX Snoep JL, Westerhoff HV;
 PI WPI; 1998-193637/17.
 XX P-PsDB; AAM56790, AAM56791, AAM56792, AAM56793.
 DR
 XX Method improving production of biomass or a desired product - by
 PT expressing an uncoupled ATPase activity in the cell
 PS Claim 16; Pages 35-41; 78pp; English.

XX This DNA encodes the soluble part (F1) of membrane bound (F0F1 type) H+
 CC - ATPase or a portion of F1 exhibiting ATPase activity. The DNA is
 CC derived from *Lactococcus lactis* subsp. *cremoris* strain MG1363. This is
 CC used in a novel method for improving the production of biomass or a
 CC desired product from a cell. The method comprises expressing an uncoupled
 CC ATPase activity in the cell to induce conversion of ATP to ADP without
 CC primary effects on other cellular metabolites or functions and incubating
 CC the cell with a suitable substrate to produce the biomass or product. The
 CC expression is directed using a vector including DNA encoding the soluble
 CC part (F1) of the membrane bound (F0F1 type) H+-ATPase or a portion of F1
 CC exhibiting ATPase activity, the DNA being derived from *Lactococcus lactis*
 CC subsp. *cremoris*, *Lactococcus lactis* subsp. *lactis*, *Streptococcus*
 CC *thermophilus*, *Plaffia rhodozyma* or *Trichoderma reesei*, where the DNA is
 CC under the control of a promoter. An ideal ATPase is the membrane bound
 CC H+ ATPase. This enzyme complex consists of two parts, the membrane
 CC integral part and the (F0) and the cytoplasmic part (F1). Together the
 CC two parts couple the hydrolysis of ATP and ADP to the translocation of
 CC protons across the cytoplasmic membrane, or vice versa. The proton
 CC gradient is used to drive ATP synthesis from ADP and Pi. The method can
 CC be used for optimising the formation of biomass or a desired product,
 CC e.g. the product may be lactic acid which results in the acidification of
 CC dairy products. The method is more efficient than currently used methods
 CC of biomass production.

XX Sequence 4815 BP; 1491 A; 897 C; 981 G; 1446 T; 0 other;

Query Match 13.28; Score 261.2; DB: 19; Length 4815;
 Best Local Similarity 50.48; Pred. No. 6.2e-55;
 Matches 1061; Conservative 0; Mismatches 838; Indels 207; Gaps 9;

QY 54 aattatcatcgtggtgcaaaaagtaactgaattatattgaacaaacccaactctccgtat 113
 DB 1967 aattgaatcgtggtcgtcgtccgttgaagtccttgaacaaacattgcacaacattg 2026
 QY 114 ctgtagcacaacaacattgatttattgtcagtagagtttggttacttaagaagattg 173
 DB 2027 ctgtgaaacaaacatttggattctctatgcatgacatcgtgcatcttgaatagttc 2086
 QY 174 actagatcgtatgtgttcatttgaatcagcaactttgagatgatacctaactatg 233
 DB 2087 cagtgtgatgtcttcttatttgaactaaatgttcgattcttctgatagcgaattatg 2146
 QY 234 ctgatttatcgtgtgagtttaaccgaatctggaattacaaatgataaataaagatgat 293
 DB 2147 cagatccttgaacgtaataactacacacaaagatttccagaagaagcaaaacttgcag 2206
 QY 294 taaagggcatttggatagcttcaagaacaaacagtgctgtgtaagttaacatttaag 353
 DB 2207 ---aagcaattaaagcattcaaaataacacgaatatttaataagaggagcttaacta 2263
 QY 354 gagagacaanaatgtaggtgtcctaagaagataagaacaaacaaatcgcgagtgttaaa 413

DB 2264 gaggcttcaact-----taacgaataaanaactaagatgcttcaacaaagaanaac 2312
 QY 414 acaaaaaattactaaagcagatggaatgtgtcgtccttgaanaatgcyftraaaacggaaga 473
 DB 2313 aagtcataacacagtgtagcagatgaaatgttctcgtcgtcaaatctcaaaagcagaaatc 2372
 QY 474 acgcagctctcttcaacgccttattcagaacaaacacagcagtaacgttattagccac---gt 530
 DB 2373 tcacgctaaagccttccagacttattcgtgaaagaagtaagtaagattacgactgactagt 2432
 QY 531 ttcaagaacacagatgtgttaccagatccatttttagtggatcgcgaagtlaaanaagt 590
 DB 2433 ttcaagcgaataatgagccgcgcgaanaaatccgatgatgattaaacgtgaagtlcaagaanaac 2492
 QY 591 gggcagatgattgtgtgtccacagatcgtgtccttctgtgtgtgttaaacgttactgtt 650
 DB 2493 tgcgtatcgttatacacaacagatcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2552
 QY 651 taaacgtgtatttaaatgaaatgaaatgaaatgaaatgaaatgaaatgaaatgaaatgaaat 710
 DB 2553 gaagctgttataagtaataatcgttatacgttatacgttatacgttatacgttatacgttatac 2612
 QY 711 ttaatcgttataagtaataatcgttatacgttatacgttatacgttatacgttatacgttatac 770
 DB 2613 tgcctgt 2672
 QY 771 agatcagatgt 830
 DB 2673 tctcgt 2732
 QY 831 gatgatgcttataaanaaggggaagtagatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 890
 DB 2733 cgtagaagaatatacagaagaagaatcgtatgaacttgaattgtgtgtgtgtgtgtgtgtgtgt 2792
 QY 891 taacacgattgtcgaagaagcagatgttgaanaaatattcattaccagaatagataa 950
 DB 2793 gaactcattgttaagtagagcagcagatgagaanaaat-----gttaccatttcttcttga 2846
 QY 951 tcatgaattagcgaagaagaacaaagtttggatattattacgaactatgsgaagaagt 1010
 DB 2847 tgaanaagtlacgaanaaagcatctctgttacaatttgaattgaacacgaatcgttgaac 2906
 QY 1011 attttagtaattattgt 1070
 DB 2907 aatcctaactcagttgttgcgcgaatattgcgaagaatgattatgttgcctcaattgttga 2966
 QY 1071 aaacttgccttgcgaagaagcgcgtcgaatgtgtgcgaatgaaagcagaacagataacgc 1130
 DB 2967 tgcanaaacagcagaacaatgtcgcagatgacgcgaatgacgtactgcgaacagataatgc 3026
 QY 1131 agttaacttaattatgattacagttacgttactatacaaaagctgttcaagaagattat 1190
 DB 3027 acattcgttcaattatgatttacaacattcaataaacytgcgtcgtcaagcttcaattac 3086
 QY 1191 aatgaatataaataattgttgcggtgcagcagc----- 1226
 DB 3087 gcaagaaatcaggaataattgttgcggtgtcctcagcgtataactacgttcaaacattat 3146
 QY 1227 ----- 1226
 DB 3147 tctcaatgttcaagttatacacttgaggaataatgttctcgtcagtaagaagcttgaat 3206
 QY 1227 ----- 1226
 DB 3207 tttaatacgttgcagtaaatlttactgtattagctttaaanaatgaaatagaaattctgt 3266
 QY 1227 -----aatttaanaaataagagatcgttgaatgcaactggaanaaatgttacaat 1276
 DB 3267 tgttagacagaanaataaacaagaaggaanaaacatttgcgttctgttataactacaggt 3326
 QY 1277 catcgtgtcgtgtatttgcgttgaatttccacaagaatgc---agtaaccaaaagtatga 1333

```

QY 1600 aatctaacgtactgaacttttagaagcgggaattaaagttatcgaacttagttgtccgt 1659
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2821 agtcacacgagatcgaaatccctcgcagacggggaataaagtatcgtcttctgcacggt 2880
QY 1660 ttgcgaagagggttaagtaggtttttcgtgtgtcgggtgtcgcgttaaacggtcaata 1719
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2881 ttccaaagggtggaaaataaggtttcttcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2940
QY 1720 ttgtgaataatccgtaacatcgcaattagacatcgaactgaactgtctgtcttgcggggg 1779
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2941 ttatggaaatttgacagaataatcgcctatagacataagggaatttccgtgttgcgggtg 3000
QY 1780 taagttagcgtacgcgttgaaggtgaactctcatatgataagtaagaaacttaacgtat 1839
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3001 ttgtgtgagagaaagaagaggaaggaacgaactctgtgttgaaatgcagaaagcggttc 3060
QY 1840 tagataaagtgtctctgtttatgtgtcaaatgaacgagccacgaagtaacggttaacgtg 1899
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3061 ttgttaacactatcctcgtcttcgtgacagatgaacgaccccggtggcaaggttcaggg 3120
QY 1500 ttgcattaacaggtcttaacatgtgcggaataattcgtga---tgaagtcgtgatgtct 1956
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3121 ttgccttgaccgcgtccacataagcgaataacttaacttaagggaggttgaaggaagagcttc 3180
QY 1957 tatcttcgttgataa 1972
    ||| ||| ||| |||
Db 3181 ttctctcatagacaa 3196

RESULT 14
AAAX14416
ID AAAX14416 standard; DNA; 1468 BP.
AC AAAX14416;
XX
XX 31-MAR-1999 (first entry)
XX H. pylori GHP0 663 gene.
XX DE
XX GHP0 protein; Helicobacter infection; gastroduodenal disease; gastritis;
XX KM peptic ulcer disease; ss.
XX OS
XX Helicobacter pylori.
XX FT
XX Key Location/Qualifiers
XX CDS 30..1439
XX FT /*tag-a
XX FT
XX MO9843478-A1.
XX PN
XX PD 08-OCT-1998.
XX PE 01-APR-1998; 98MO-US06371.
XX PR 29-JUL-1997; 97US-0902615.
XX PR 01-APR-1997; 97US-0833457.
XX PR 24-JUN-1997; 97US-0881227.
XX PA (HUMA-) HUMAN GENOME SCI INC.
XX PA (INMR) MERIEUX ORAVAX PASTEUR MERIEUX SERUMS.
XX PI Al-Garawi A, Kleanthous H, Miller C, Oomen RP, Tomb J;
XX DR WPI: 1998-542293/46.
XX DR P-PSDB; AAW98697.
XX
XX New isolated Helicobacter polynucleotides - used to develop products
XX PT for the diagnosis, prevention and treatment of Helicobacter
XX PT infections and gastrointestinal diseases
XX PS
XX Claim 1; Page 1456-1458; 2054pp; English.

```

```

CC This sequence represents a polynucleotide of the invention. It was
CC isolated from Helicobacter pylori and encodes a H. pylori GHP0 protein.
CC The polypeptides can be used for preventing or treating Helicobacter
CC infections, and gastroduodenal diseases associated with these
CC infections, including acute, chronic, and atrophic gastritis, and peptic
CC ulcer diseases, e.g. gastric and duodenal ulcers. They can also be used
CC for the production of antibodies. The products can also be used for
CC detection and diagnosis.
XX
XX Sequence 1468 BP; 418 A; 248 C; 404 G; 398 T; 0 other;
XX
Query Match 10.9%; Score 215.4; DB 19; Length 1468;
Best Local Similarity 58.8%; Pred No. 8.6e-44;
Matches 372; Conservative 0; Mismatches 261; Indels 0; Gaps 0;
QY 1311 gatcgatcaacaagatataatgtacgttaacttaattgttaaacaggatttagtactgaagt 1370
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 114 gaagcgttagacatttaatttgaagtaatggtgttcaaaagctttagtttagaggtg 173
QY 1371 caacaacaaatgaagtggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1430
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 174 gcagccattggcggttaatcggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 233
QY 1431 aaacgcggtttagcggtgaacaaatgaacaaattcttccagtggtgaacga 1490
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 234 gtgtgttaaccaagtgatcaaggtcgcgcgcaaaatgttaagtgtcgtgtgtgtgtgtgtgt 293
QY 1491 acattggtctgatcatgaacgtatgtgtgaaccaaactcatgtagcaaggtgaatcgt 1550
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 294 gtattagggcgtattttaaattgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 353
QY 1551 gcagaagagaattgttcatcaccgtgtgcgcacaaagttaagaagaacaatacga 1610
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 354 ccgtcctaacttggccattccacagaagcccttagttttagcagcaagcactaa 413
QY 1611 actgaacttttagaagaagggaatgaagttatgcacttagtttgcgtgtgcgaagg 1670
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 414 acagaatgtttgaactgtatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatga 473
QY 1671 gtaagaatgattatcctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1730
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 474 ggtaaagttagctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 533
QY 1731 atccgttaacatgcgaatttgacactcaggttaactgtcttgcgggtgtgtgtgtgtgtgt 1790
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 534 atccataatgtgtttaaagcataacaggtattcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 593
QY 1791 acgcgtgaagtaacgacttcatcatgagatgaagaagtaacacttaacgtattagataaagt 1850
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 594 accagagagggaatgtatctgtatttgaatgaagaagaagggtgttttagacaaagtc 653
QY 1851 tctctgtttatgttcaaatgaacgagccacaggttaacggtttagctgtgtgtgtgtgtgtgt 1910
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 654 gcactgtgtttatgtgcaaatgaatgaatgagccacaggtgtgtgtgtgtgtgtgtgtgtgtgt 713
QY 1911 ggtctaactatggcggaaaatccgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1943
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 714 ggttgacgtatgtgcgaggtatttctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 746

RESULT 15
AAZ32024
ID AAZ32024 standard; DNA; 8008 BP.
AC AAZ32024;
XX
XX 10-JAN-2000 (first entry)
XX DE
XX Human METH1 related EST X99599.
XX DE
XX Human METH1; METH2; anti-angiogenic; metalloproteinase thrombospondin;
XX KM cancer; diagnosis; hyperproliferative disorder; autoimmune disease;

```

```
Db 3327 tatcgctcccgctggtgacgtggaatttggtctgtatgccaactgcctgagattacaac 3386
QY 1334 tgccttaaty-----tgaacaaggtttactgtaagttca 1372
FT ||||| | |
Db 3387 tgccttgattgtctacaagaatgtcaatgtttaaaaaaataactcttgaagttgc 3446
QY 1373 acaacattagggggggtggttgctgcgtatcgcaatgggggtcatctgattgattaa 1432
Db 3447 tttagaacttggtagtgcagctctgacgactgataaggaactcaactgattgattgac 3506
QY 1433 acgcgatttaagcgtacaataatgcataacccaattctgtccagtgggaacgaaac 1492
Db 3507 tctgtgacttgaagtccttgatatacaggttaaggggtccagcgttccctgttggtaactac 3566
QY 1493 attggctgatalcaatgaacgtatctgggtgaaccaatcgatbgcaagggtgaaatcggtgc 1552
Db 3567 tcttgctcggttcttaactgctccttggtgacgttattgattggtgggaagatttccctgc 3626
QY 1553 aga---aggaattggtctatccaccgtgcgcacaaagttaagaagaacaattacaag 1609
Db 3627 tgatgagaacgtaactccatccacaagaagaagcttcaactttgacgaaattgtcaactgc 3686
QY 1610 tactgaactttagaacgcggaaatlaaagttaacgttaactgaatttgcggtttgcgaag 1669
Db 3687 aatatgaagtctctgtacaaggagtaaatgttgcgatttcaactgccccttactttaagg 3746
QY 1670 gggtaaatgagtttacttcgtggtgcgggtgcgtgtaaaacgcgtcaataatgataagtt 1729
Db 3747 tgggaagtcgcgactcttcgtggtgcgggtgtgtaaaacgcgtctcttccaagaatt 3806
QY 1730 aatcgtacaatcgcaattgagagcactcaagttactgtcttctggcggggtaaggtaagc 1789
Db 3807 gatccaataatgccaagaacacggtgtatcttctgtattcaacaggtgttgcgacgc 3866
QY 1790 tacgcgtgaagtaacgacttctatcatgagatgaagaagcttaacgttaataaagt 1849
Db 3867 tactcgtgcgggaatgactcttactcgtggaagaagaatacgcggttatttgaaaaaaac 3926
QY 1850 gtctcttgttatgtgtcaaatgaacgagccacaggttaacggttctgtgtgcatlaac 1909
Db 3927 agccatggtcttgtgtcaaatgaatgaacacactgagacggtatgctgttgccttac 3986
QY 1910 aggccttaactatgacggaanaatcgcgtatgt---aagtcgtgagtcttacttctcgt 1966
Db 3987 tggtttaacaatgcggaatattccgtgtatgttcaaggacaagacgtatttcttcat 4046
QY 1967 tgataa 1972
Db 4047 cgataa 4052

RESULT 13
AAK29857
ID AAK29857 standard; DNA; 4001 BP.
AC
AC AAK29857;
XX
XX 22-JUN-1999 (first entry)
DT
XX
XX Thermotoga neapolitana type F ATPase gene.
DE
XX
XX Heat resistant; type F ATPase; enzyme; bioreactor; ds.
KW
XX Thermotoga neapolitana.
OS
XX
XX Key Location/Qualifiers
FH 130..1641
FT CDS
FT /tag= a
FT /product= "subunit #1"
FT /transl_except= (pos: 130..132, aa: Met)
FT CDS
FT /tag= b
FT /product= "subunit #2"
```

```
FT /note= "CDS appears to be missing 48 bases between
FT nucleotides 2040-2041"
FT CDS 2448..3806
FT /tag= C
FT /product= "subunit #3"
FT /note= "CDS appears to be missing 48 bases between
FT nucleotide 3386-3387"
PN JP11075855-A.
XX
XX 23-MAR-1999.
XX
XX 16-SEP-1997; 97JP-0250926.
XX
XX 16-SEP-1997; 97JP-0250926.
XX
XX (KAIYO) KAIYO BIOTECHNOLOGY KENKYUSHO KK.
DR WPI: 1999-257703/22.
DR P-PSDB; AAY02647, AAY02648, AAY02649.
XX
XX Heat resistant type F ATPase - useful as a heat resistant reagent
PT and an ATP-supplying enzyme in a bioreactor
XX
XX Disclosure; Page 9-13; 14pp; Japanese.
XX
XX This sequence represents a novel Thermotoga neapolitana heat resistant
CC type F ATPase gene which encodes 3 subunits of the protein complex.
CC The new protein is useful as a heat resistant reagent and an
CC ATP-supplying enzyme in a bioreactor.
XX
XX Sequence 4001 BP; 1185 A; 872 C; 1095 G; 849 T; 0 other;
SQ

Query Match 12.7%; Score 250.8; DB 20; Length 4001;
Best local Similarity 57.6%; Pred. No. 2,2e-52;
Matches 528; Conservative 0; Mismatches 352; Indels 36; Gaps 3;

QY 1093 gctcgaatgtgcgcattgaagaagcaacagataacgcaggttaacttaattgattga 1152
Db 2281 gccagacagaaccccttgagaagacgcaacgcagcaagcaggaagtgaacagaagctg 2340
QY 1153 cagttgcttatatacaaatcgtcgaagcaagatattacaatgaattgaattgtt 1212
Db 2341 acactggtcttacaacaagaacgctcaggtcgtgatcacccaggaactcatagagatgt 2400
QY 1213 gccggtgcacgacgaatttaacaatalagagatcggtaatg----- 1254
Db 2401 acggtgtcgaagcaactggaaggaatcgaaaatgagtgatagcgtatgcaaaaggtt 2460
QY 1255 caactggaaaaattgtacaatcatcgtgtgcgtttatgtacgttgaattccacaagatg 1314
Db 2461 caaaagatacatagtggaagtgatggtggtccgtcgttgatgtgaagttcccgagaag 2520
QY 1315 cagtaccaaaagtatagtatgcttaaatgttgaacaggtta-----g 1359
Db 2521 aacttccaacatatcaacgaccttgaggtcgtgatcatcctcaaaaggggcagaagattg 2580
QY 1360 tacttgaattcaacaacaattagtggtgtgtgtatgtcgtgtalacgaaatggatcat 1419
Db 2581 ttcttgaagtagaacaaccatagagacgtgtcgtggaacggtggtgatgtactcca 2640
QY 1420 ctgattgattaaacggtttaaagcgtatacaaaatcgataaaccatttctgtccag 1479
Db 2641 ccgattggtcatalgaagggcttgaagtggttgaatcgggttgaaccaattacgcgtcccg 2700
QY 1480 tgggaacgaaaacatggtgtcgtatcatgaacgtattggttggaaccaatcgatgaacag 1539
Db 2701 tgggaagaagaatcaccggaagaatactgaacgttataagagagccgttggagaagccg 2760
QY 1540 gtgaatcgtgtcagaagaatgtgtatcatccacgtgcgcacccaagttaagaagac 1599
Db 2761 gcgagataaagtcogaaggaaggttggcgcgtacacagggcgtcctcgtgaactatagaac 2820
```

KM angiogenesis inhibitor; abnormal wound healing; inflammation;
 KM rheumatoid arthritis; psoriasis; endometrial bleeding disorder;
 KM diabetic retinopathy; macula degeneration; haemangioma; detection;
 KM arterial-venous malformation; immune deficiency; ss.
 OS Homo sapiens.
 PN W0937660-A1.
 PD 29-JUL-1999.
 PF 22-JAN-1999; 99WO-0501313.
 PR 23-JAN-1998; 98US-0072298.
 PR 28-AUG-1998; 98US-0098539.
 XX (IRUE/) IRUELA-ARISPE L.
 PA (HAST/) HASTINGS G A.
 PA (RUBE/) RUBEN S M.
 PR Irue-la-Arispe L, Hastings GA, Ruben SM.
 DR WPI: 1999-590684/50.
 PT New isolated metalloprotease thrombospondin polypeptides, useful for
 PT treating hyperproliferative disorders, cancers or autoimmune disorders
 PT
 XX Disclosure; Page 347-352; 457pp; English.
 PS
 CC AA232000 and AA232001 encode, and AA49501 and AA49502 represent, human
 CC metalloprotease thrombospondin (METH) proteins METH1 and METH2
 CC respectively. METH1 and METH2 have been found to be potent inhibitors of
 CC angiogenesis both in vitro and in vivo. They can be used for treating
 CC cancer and other disorders related to angiogenesis including: abnormal
 CC wound healing, inflammation, rheumatoid arthritis, psoriasis,
 CC endometrial bleeding disorders, diabetic retinopathy, some forms of
 CC macula degeneration, hemangiomas, and arterial-venous malformations.
 CC They may be useful in treating deficiencies or disorders of the immune
 CC system, by activating or inhibiting the proliferation, differentiation,
 CC or mobilization (chemotaxis) of immune cells. The etiology of these
 CC immune deficiencies or disorders may be genetic, somatic, such as
 CC cancer or some autoimmune disorders, acquired (e.g. by chemotherapy or
 CC toxins), or infectious. They can also be used to treat inflammatory or
 CC conditions, both chronic and acute conditions. The products can also be
 CC used for detection and diagnosis. AA232002 to AA232080, and AA49503 to
 CC AA49511 represent sequences given in the exemplification of the present
 CC invention.
 XX
 XX Sequence 8008 BP; 1430 A; 2591 C; 2649 G; 1338 T; 0 other;
 SQ
 Query Match 10.6%; Score 210; DB 20; Length 8008;
 Best Local Similarity 53.6%; Pred. No. 3.3e-42;
 Matches 548; Conservative 0; Mismatches 440; Indels 34; Gaps 4;

Db 5490 cgaagcgtctgacggaacccgagatagaagaatggaagcaagcaagcaagtgaccga 5549
 QY 1274 aatcatcggtggtgtatgtacgttggaatcccaaga-----tgaatgcaaaagt 1327
 Db 5550 ggtcatcggtcggttcgtacgtgagtcgaagaagcgtcccggtatctgaacgc 5609
 QY 1328 atatgagccttaaatgttgaacaggttgaactgaactgaactgaactgaactgaact 1387
 Db 5610 ccttgaacaccacaacaagcaagcagcgtcttctcgaagtggtggtggtggtggtg 5669
 QY 1388 tgggtgagtcgtgtatgcgaatggatcgtatgatgataaagcgttgaagcgt 1447
 Db 5670 gaacaccgtcgacacatcgatgacgagcagaggtctcgtggtggtggtggtggt 5729
 QY 1448 aacaaatgaatgaacccaattctgttcgaatgggaagaaacattggtggtatcat 1507
 Db 5730 gtcgacaccggtggtggtggtggtggtggtggtggtggtggtggtggtggtggt 5789
 QY 1508 gaacgtattggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtg 1567
 Db 5790 gaacgtatcggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtg 5849
 QY 1568 tatcaacgtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtg 1627
 Db 5850 gatcaacacccggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtg 5909
 QY 1628 ggaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatga 1687
 Db 5910 cggcatcaagtgatgacgtggtggtggtggtggtggtggtggtggtggtggtggtg 5969
 QY 1688 cgtgtgtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtg 1747
 Db 5970 cgggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtg 6029
 QY 1748 tgaacgtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtg 1807
 Db 6030 agtgcactggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtg 6089
 QY 1808 ctctatcatgagatgaagactgaactgaactgaactgaactgaactgaactgaact 1846
 Db 6090 ccttaccacagatgatgatgatgatgatgatgatgatgatgatgatgatgatgat 6149
 QY 1847 agtgcactggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtg 1906
 Db 6150 agtgcactggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtg 6209
 QY 1907 aacaggttaactatggtggtggtggtggtggtggtggtggtggtggtggtggtggt 1966
 Db 6210 gacgtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtg 6269
 QY 1967 tg 1968
 Db 6270 cg 6271

Search completed: October 28, 2001, 17:20:40
 Job time: 8409 sec

QY 683 aagaaaaagatgttccgttcaattgaatttaacgtgttcaatctatcaacttttcc 742
 DB 7454 AAGAAAAAGATGTTTGTGTTCAATGATTAATCGTTCTAAATCATATCACTTTTCC 7513
 QY 743 aactcttggaattaaatttaacccaagattccagatttggtgtaactccctctgtg 802
 DB 7514 AATCTTGGGAAATTAATTAATTAACCAAGATTGAGTATTTGGTGATACCTCCCTGTTG 7573
 QY 803 agcagtaattgtgtcagtlcaattctatgattgagcttaaaaaaggggaagtagatg 862
 DB 7574 AGGAGTTAATTTGGTTCAGTCAATTTCTATGATGATCTTTAAAAAGGGAAGTAGATG 7633
 QY 863 ttgtgtattagttataaacaatttataacagatgttcgcaaaagccggtttgaaa 922
 DB 7634 TTTGTGTTAGTTTATTAACAAATTTATTAACAGATGTGCAAAAGCCAGTATTTGAAA 7693
 QY 923 aattaatccattaccaggaattgatatatgaattaggcgaaagaacaaagcttcg99 982
 DB 7694 AATTAATTCATTAACGAATTAATGATTAATGATTAATGAGCGAAAGAAACAAAGTTTGGG 7753
 QY 983 attataattcgaacctgagtcgaaagatatatagataatttatttggttcgtatttag 1042
 DB 7754 ATTTATTTTACGAACCTGATGCGGAATATTATTAATTAATTTATTTGCTTATTTAG 7813
 QY 1043 aatctcaggtttatcaagcagcagttgaaaaacttgccttcgtgcaagccgctcgaaatg 1102
 DB 7814 AATCTCAGTTTATCAAGCAGCAAGTGAACCTTGTCTTGACCAAGCCCTGGAATGG 7873
 QY 1103 tcgcaggaagaagcgcaagagataagcaggaatttaattgaattgaattgaattgct 1162
 DB 7874 TCGGCAATGAAGAGCGCAACATTAACCGAGGTAACTTATTAATGAGTTACAGTTAGCT 7933
 QY 1163 ataacaagctcgtcaagcaagtaattacaattgaattgaattgttgcggttcag 1222
 DB 7934 ATTAACAAGCTCGTCAAGCAAGTATTAACAATTAATTAATTAATTTGCTCGGTGAG 7993
 QY 1223 cagcaatttaacaatgagagatcgtgtaattgcaacttggaataattgtaacaatcatctg 1282
 DB 7994 CAGCAATTTAAACAATGAGATCGTAAATGCAACTGGAATAATTTGCAAAATCATGG 8053
 QY 1283 tgcggttattgaattgaattcccaagaagtcagtaacaaagatatatgcttcctaaa 1342
 DB 8054 TCGCGTTATTTACCTTGATTTCCCAAGATGAGTACCAAAAGTATGATGCTTAAA 8113
 QY 1343 tgttgaacaaggtttagtacttgaagttcaacaacaattgagtggtgtgtagtgcctg 1402
 DB 8114 TGTGAACAGGTTTAGTGTGAAGTTCAACAACAATTAATTAATTTAGTGTGTGAGTGTG 8173
 QY 1403 tatcgcaatgggattcatctgtgatttaaaocggtttaagcgtaacaataagcaataa 1462
 DB 8174 TATCGCAATGGGATCATCTGTGATGATTAACGCGGTTTAAGCGTAACAAATACGAATTA 8233
 QY 1463 cccaatttcgttcagtgagggaacgaacaacatttggtcgtatcatgaaacgtatgtgtga 1522
 DB 8234 CCCAATTTCTGTCCACAGTGGGAACCAAAACAATTTGGGTGTATCAAGCAAGTATTTGGGGA 8293
 QY 1523 accaatcgaagcaaggttgaatcgtgtgcagaagaagaattggtctatctaccgtgcgc 1582
 DB 8294 ACCAATTCGATGAGCAAGGTGAATGCGTGTGCAAGAAAGAAATTTGCTATTCACCGTGC 8353
 QY 1583 accaagttatgagaagaacatctaaacagttactgaacttttagaagaagggaattaaagttat 1642
 DB 8354 ACCAAGTTATGGAAGAACATCTAACAGTACGTAACCTTTTGAAGCGGGAATTAAGTTAT 8413
 QY 1643 cgaactagttgttcggtttgcgaagaagggttaagtagttatcttcggtgtgcggtgtg 1702
 DB 8414 CGACTTAGTTTGTCCGTTTGCAGAAAGGGGTTAAAGTATGTTATTCGTTGCGGGGTGT 8473
 QY 1703 cgttaaacacgtcaaalatgataaggaattaaacgttaacatcgcaattgagactcaagttta 1762
 DB 8474 TGGTTAAACCGTCAATATGATGGAATTAATTCGTTAAACATCGCAATTTGAGCACTCAGTTA 8533

QY 1763 ctctgtcttgcgggggtaggtgagcgtgagcgtgaaagtaacgactttcatatgagat 1822
 DB 8534 CTCTGTCTTTGCCGGGGTAGTGAGCGTACGCGTGAAGGTAAAGCACTTATATATGAGAT 8593
 QY 1823 gaaagactcctaagttatgataaagtgctctgttttattgagtcgaatgaagagccacc 1882
 DB 8594 GAAAGACTCTTAACGTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 8653
 QY 1883 aggttaacgctttagcgtgtgcatcaacagccttaactatgagcggaaataattcgtgatga 1942
 DB 8654 AGGTAAACGTTTACGTTGTGCTGATTAACAGGCTTAACCATGCGGAAATAATTCGTGATGA 8713
 QY 1943 aggtcgtgactgtcttattcttcgttgataa 1972
 DB 8714 AGGTGCTGATGCTTATTTCTGTTGATTA 8743
 RESULT 2
 U32730 13828 bp DNA BCT 29-MAY-1998
 LOCUS Haemophilus influenzae Rd section 45 of 163 of the complete genome.
 DEFINITION U32730 L42023
 ACCESSION U32730.1 GI:3212191
 VERSION U32730.1 GI:3212191
 KEYWORDS
 SOURCE Haemophilus influenzae Rd.
 ORGANISM Haemophilus influenzae Rd.
 Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
 Haemophilus
 1 (bases 1 to 13828)
 Fleischmann, R.D., Adams, M.D., White, O., Clayton, R.A.,
 Kirkness, E.F., Kerlavage, A.R., Bult, C.J., Tomb, J., Dougherty, B.A.,
 Merrick, J.M., McKenney, K., Sutton, G.G., Fitzhugh, W., Fields, C.A.,
 Gocayne, J.D., Scott, J.D., Shirley, R., Liu, L.I., Glodek, A.,
 Kelley, J.M., Weidman, J.F., Phillips, C.A., Spriggs, T., Hedblom, E.,
 Cotton, M.D., Utterback, T., Hanna, M.C., Nguyen, D.T., Saudek, D.M.,
 Brandon, R.C., Fyne, L.D., Fritchman, J.L., Fritchman, J.L.,
 Georgiagen, N.S., Gnehm, C.L., McDonald, L.A., Small, K.V., Fraser, C.M.,
 Smith, H.O. and Venter, J.C.
 Whole-genome random sequencing and assembly of Haemophilus
 influenzae Rd
 Science 269 (5223), 496-512 (1995)
 MEDLINE 9350630
 REFERENCE 2 (bases 1 to 13828)
 AUTHORS Tatusov, R.L., Mushgenian, A.R., Bork, P., Brown, N.P., Hayes, W.S.,
 Borodovsky, M., Rudd, K.E. and Koonin, E.V.
 Metabolism and evolution of Haemophilus influenzae deduced from a
 whole-genome comparison with Escherichia coli
 Curr. Biol. 6 (3), 279-291 (1996)
 TITLE
 JOURNAL
 MEDLINE 96398784
 REFERENCE 3 (bases 1 to 13828)
 AUTHORS White, O., Clayton, R.A., Kerlavage, A.R. and Fleischmann, R.D.
 Direct Submission
 Submitted (25-JUL-1995) The Institute for Genomic Research, 9712
 Medical Center Dr, Rockville, MD 20850, USA
 REFERENCE 4 (bases 1 to 13828)
 AUTHORS White, O., Clayton, R.A., Kerlavage, A.R. and Fleischmann, R.D.
 Direct Submission
 Submitted (27-SEP-1997) The Institute for Genomic Research, 9712
 Medical Center Dr, Rockville, MD 20850, USA
 The H. influenzae sequence has been updated by R. Fleischmann. New
 database matches have been assigned, product names have been
 improved, and a number of frame shifts have been corrected. We
 gratefully acknowledge the work of Tatusov et. al. We have
 incorporated their annotation into the /notes fields of the
 corresponding H. influenzae genes
 5 (bases 1 to 13828)
 REFERENCE White, O., Clayton, R.A., Kerlavage, A.R., Fleischmann, R.D.,
 Peterson, J., Hickey, E., Dodson, R. and Gwinn, M.
 Direct Submission
 Submitted (28-MAY-1998) The Institute for Genomic Research, 9712
 Medical Center Dr, Rockville, MD 20850, USA
 The whole genome was shifted by 588 nucleotides for a new start
 On Jun 12, 1998 this sequence version replaced gi:1573448.
 COMMENT

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 28, 2001, 14:58:41 ; Search time 2621.89 Seconds
(without alignments)
11633.746 Million cell updates/sec

Title: US-09-545-199C-3
Perfect score: 1972
Sequence: 1 agcgggccattgctcagf.....gtctattctctgtgataa 1972

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapept 1.0

Searched: 1344157 seqs, 7733874588 residues
Total number of hits satisfying chosen parameters: 2688314

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl: *
1: gb_ba1: *
2: gb_ba2: *
3: gb_ba3: *
4: gb_in1: *
5: gb_in2: *
6: gb_in3: *
7: gb_om: *
8: gb_ov: *
9: gb_pat1: *
10: gb_pat2: *
11: gb_ph: *
12: gb_pl1: *
13: gb_pl2: *
14: gb_pl3: *
15: gb_pl4: *
16: em_ba1: *
17: em_ba2: *
18: em_fun: *
19: em_htgo_hum: *
20: em_htgo_inv: *
21: em_htgo_rtd: *
22: em_htg_hum1: *
23: em_htg_hum2: *
24: em_htg_hum3: *
25: em_htg_hum4: *
26: em_htg_hum5: *
27: em_htg_hum6: *
28: em_htg_hum7: *
29: em_htg_hum8: *
30: em_htg_inv1: *
31: em_htg_inv2: *
32: em_htg_other: *
33: em_htg_rtd: *
34: em_hum1: *
35: em_hum2: *
36: em_hum3: *
37: em_hum4: *
38: em_hum5: *
39: em_hum6: *
40: em_hum7: *
41: em_in: *
42: em_om: *
43: em_or: *

44: em_ov: *
45: em_pat: *
46: em_ph: *
47: em_pl: *
48: em_rtd: *
49: em_sts: *
50: em_sy: *
51: em_un: *
52: em_v1: *
53: gb_sts1: *
54: gb_sts2: *
55: gb_sts3: *
56: gb_sy: *
57: gb_un: *
58: gb_v11: *
59: gb_v12: *
60: gb_htg1: *
61: gb_htg2: *
62: gb_htg3: *
63: gb_htg4: *
64: gb_htg5: *
65: gb_htg6: *
66: gb_htg7: *
67: gb_htg8: *
68: gb_htg9: *
69: gb_htg10: *
70: gb_htg11: *
71: gb_htg12: *
72: gb_htg13: *
73: gb_htg14: *
74: gb_htg15: *
75: gb_htg16: *
76: gb_htg17: *
77: gb_htg18: *
78: gb_htg19: *
79: gb_htg20: *
80: gb_htg21: *
81: gb_htg22: *
82: gb_htg23: *
83: gb_htg24: *
84: gb_htg25: *
85: gb_pr1: *
86: gb_pr2: *
87: gb_pr3: *
88: gb_pr4: *
89: gb_pr5: *
90: gb_pr6: *
91: gb_pr7: *
92: gb_pr8: *
93: gb_pr9: *
94: gb_fo1: *
95: gb_fo2: *
96: gb_in4: *
97: gb_pr10: *
98: em_ba3: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match length	ID	Description
1	1925.2	97.6	10229 1	AE006185 Pasteurel
2	1255.6	63.7	13828 3	U32730 Haemophilus
3	870	44.1	870 2	AF237922 Pasteurel
4	812	71.1	2	ECOHATP W2546 E.coli H+ A
5	812	41.2	11414 1	AE000450 Escherich
6	812	41.2	14526 2	ECUNC X01631 E. coli ori
7	812	41.2	136254 2	ECOWB2 L10328 E. coli; th
8	810.4	41.1	11999 1	AE005605 Escherich

David Lavery
Tony E. Guller
Michael J. Kennedy

David Lavery
Tony E. Guller
Michael J. Kennedy
April 6, 2002

[illegible][illegible]

```

/codon_start=1
/transl_table=1
/product="AtpE"
/protein_id="AAK03573.1"
/db_xref="GI:12721873"
/translation="METVITTTTIFASAILLAVALGATGALGKFLSSAROE
ASSQIKMFIVAGLDALSMIVAGIALLFIFANPFDLLK"

```

/transl_table=11
/product_id="Atpa"
/db_xref="GI:12721876.1"
/translation="MQLNSTEISELIKRIADFDVSEARNTGTVSVSDGIIRHGL
SEVMOGEMIALPGRNVALNLEEDSVAVWGPADIAEGMEVCTRIIEVPGRL
LGRVNTLGEIDGKEIDENDGSPVAVIADRSVDPVOTGYKAVASMPIG
RGORELIGDROTKTALAIIINORSGIKCIYVAIGKASTIANVVRKEEHSAL
ONTIVVASASESALQYLAISYSCAMEPDRREDLIIYDDSKAVAYROSL
LRRPGRAPFDVFTYLSRLERASRNEEVEKFTNGEVKGTGSLALPIETQA
GDVAFVPTNVIISTDQIFLESNLNSGIRPAVPGISVSVGSSAQTKYKLAG
IRTAQYRELAAFAQFASDIDATRKQISGOKVTELSKQYSPLSVAQALFLA
VEFGYLEVEDLDRIQSFESALLEYANHNVADEMRLETQSGNNDIEIKESLGIIDSPK
ANSAW"
gene 7135. 8004
/gene="atpc"
/note="PM1493"
7135. 8004
/gene="atpc"
/codon_start=1
/transl_table=11
/product="Atpc"
/protein_id="AAK03577.1"
/db_xref="GI:12721877"
/translation="MAGAKERTKIASVSTOKITKAMENYVAAKSRKTOERMSSRP
YSEIRNVIHVSKATIGKHPFLVDREVKKVMIVVSTDRGLCGGLVNVLEKTYLNE
MKEKREKVSQSLISGKSTNFRSLIKILITDSDGIDTPSYQLIGVNSMIDAY
KKEVDVYLYNKTINRMSQRPVLEKLIPLPELDNDLGEKQVMDIYEDDAVLL
DNLRIYLESQVQAAVENLASDAQARVAMKAANDNNGNLINELQVYNNKARQASIT
NELNEIVAGAAAI"
gene 8022. 9395
/gene="atpd"
/note="PM1494"
8022. 9395
/gene="atpd"
/codon_start=1
/transl_table=11
/product="Atpd"
/protein_id="AAK03578.1"
/db_xref="GI:12721878"
/translation="MATGRIVQIIGAVIDVEFPODAVRKYVDALNVTGLVLEVOQL
GGVVRCLIAMGSDGLKRLSVNTNTNISVPVGTGRIMNVNLEGEIDEGEIGAE
ENMSIHRAPSPYEQSNSTELLENGIKYIDVCPAKGKGLFGAGVCGKVNAAE
IRNIAIEHSQSVFAGVEERTREGNDFHEMKDSVYDKVSLVYGOMEPNCRNRA
LGTMAERDEGRDVI.FYDNIYRTILACTEVSALLGRMPSANGYPTLAEEGVL
QERTSTKGSITSQAVYVADDTLDSPTATFNHDSIVVLSQISLIGLYPVRD
LESTSRDPLVGEENHVARGVQTLQARKELDIILIGMELSEDKIVARAR
KIERFLSQPFVAEYFNCTPKYVPLKETIRGFKGILDEYDHIPEQAFYMAGTIDEV
LEKAKKL"
gene 9436. 9864
/gene="atpc"
/note="PM1495"
9436. 9864
/gene="atpc"
/codon_start=1
/transl_table=11
/product="Atpc"
/protein_id="AAK03579.1"
/db_xref="GI:12721879"
/translation="MSVFNLTIVVSAEQIFSGOVESIOATGIEGLIAGHTPLTA
IKPISVLTLEDGKEVYVSGFLEQPNVTVIADIRGDELQDGRILAAKKRAE
ENIRTRGDANYEMLASKEELAKIRAYELTEKLIVKKR"
gene complement(9922. 10161)
/gene="PM1496"
CDS complement(9922. 10161)
Query Match 84.9% Score 1674: DB 1: Length 10229;
Best Local Similarity 99.7% Pred. No. 0;
Matches 1924; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 104 tctccgtatctgtagcacacaagcattagtgatttgcagtagagttgtagtacta 163
DB 6875 TCTCCGTTATCTGTACACACAGCATTAGTGTATTTGCGATGAGTTGGTTACTTA 6934
QY 164 gaagaagtgcactatgattgatttgcatttgcacacacatttgaagtgcatac 223
DB 6935 GAAGAAGTGCAGCTTACATGCTATGTTGTTGATTCAGCACTTTTACAGTATGCTAAC 6994
QY 224 cataactatgcatttatttgcgtgagtttaaccacacatcgtcaattacagtgaatt 283
DB 6995 CATTAAGTGCAGCTTATTTGCTGAGTTAACCCATCTGCGCAATTCAGTATGAGTAAT 7054
QY 284 aaagagcttaaaagcatttgcatttgcatttgcatttgcatttgcatttgcatttgc 343
DB 7055 AAAGAGCTTAAAGCATTGCTTGTGATCTTCAAGCAAGCTGCTGCTGATGATTAAC 7114
QY 344 actttaaagagagacaaatgagcagtgctaaagaagataaagaacaaatcgcagtg 403
DB 7115 ACTTTAAATGAGAGACAAATATGCGAGTGTAAAGATTAAGTAACCAATTCGAGTG 7174
QY 404 taaaagtaacacaaaattactaaagcagtggaatggtgctgctcgcgaatgagta 463
DB 7175 TAAAAGTACACAAAATAATTACTTAAAGCATGGAATGCTGCTGCTGCAAAATGCGTA 7234
QY 464 aaagcagaagacagcagtccttctcagccctatttcaagaacaaatcagtaacgtgatta 523
DB 7235 AAAGCAGAAGCAGATGCTTCTTCAAGCCCTTATTCAGAAACAAATACGTAACGTATTA 7294
QY 524 gcaagcttccaaagcaagcattggttacaagcattcattttagtgcagta 583
DB 7295 GCCAGCTTCCAAAGCAGCAGATTGTTACAGATGATTTTATGAGTACCGGAGGTA 7354
QY 584 aaaaagtgagcagtgatgctgctcagcagatcgtgcttctgagtgagtgagtgagtgag 643
DB 7355 AAAAAGTGGCAGTATGTTGTTGCTCCAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 7414
QY 644 acttggttaaaactgatttaaaatgaatgaatgaatgaatgaatgaatgaatgaatgaatga 703
DB 7415 ACTTGTTAAACTGATTAAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGA 7474
QY 704 aattggtttaaactgcttcaaatcacaacttccaaacttccaaacttccaaacttccaaacttcc 763
DB 7475 AATTGAGTTAATCGGTCTTAAATCTATCACTTTTCCAACTTTTGGGAATTTAAATTT 7534
QY 764 taaccacagttcaggtatggtgagatccctcgttgcagcagttatggttgcagta 823
DB 7535 TAACCAAGATTCAAGTATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 7594
QY 824 attctatgattgattcttataaaagagagagtagatggttgcattatgattataca 883
DB 7595 AATTGATGATTGATGCTTATTAATAAAGGAGGAGTATGATGATGATGATGATGATGATGATGATG 7654
QY 884 aatttatacagcagatggtcgaagaagcagatttgcgaagaatatttccattaccagaat 943
DB 7655 AATTATTTAACAGATGCTGCAAAAGCCAGTATGCAAAATTAATTCATTAACAGAAAT 7714
QY 944 tagataatgatgaattagcgaagaagaagaagttgagattatattcgaacattgag 1003
DB 7715 TAGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 7774
QY 1004 ggaagatattatgaataattatggttgcattatggaatcgaattcgaagcag 1063
DB 7775 CGAAAGTATTTAGATTTATTTATGCTGCTGATTTAAGATTCAGGTTATCAAGCAG 7834
QY 1064 cagttgaaacattgcttgcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1123
DB 7835 CAGTTGAAACCTTGCCTTGTGAGCAAGCCGCTGAGTATGCTGCTGAGTATGAGTATGAGTATGAG 7894
QY 1124 ataacgagtaacttaataatgagttcagttgattcagatacaagaagctgcagca 1183
DB 7895 ATTAAGCAGGTAAGTATTAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 7954

[illegible]

JOURNAL	Unpublished
REFERENCE	2 (bases 1 to 870)
AUTHORS	Fuller,T.E., Kennedy,M.J. and Lowery,D.E.
TITLE	Direct Submission
JOURNAL	Submitted (24-FEB-2000) Discovery Research, Pharmacia & Upjohn Animal Health, 7923-25-434, 7000 Portage Road, Kalamazoo, MI 49001-0199, USA
FEATURES	Location/Qualifiers
source	1..870
gene	/organism="Pasteurella multocida"
	/db_xref="taxon:747"
CDS	1..870
	/gene="atpG"
	1..870
	/gene="atpG"
	/codon_start=1
	/translation="1
	/product="ATP synthase F1 gamma chain"
	/protein_id="AAF68408.1"
	/db_xref="GI:7716505"
	/translation="MACAKEIRTKIASVKSHTOKITKAMENVAASKRKTDERMSSSRP
	YSEPIRNVISHVSKATIGYKHPFLVDEVRKVGAVIVSTDGLGGLNVLFKVLNE
	MKEKREKDVSVOLSLISKSTINEROSIGIKILNDDSGIGDPVEOLIGSVNSIDA
	KKGVDDVYLVYTNKFIINTMSOKPYLEKIPIPELDNDEBGRKQVMPYIYEPDAKVLII
	DNLIVRYLESQYVQAAVENLASSEQAARVAVAKATDNGINLNEQLYVINKARASIT
	NELNEIVAGAAI"
BASE COUNT	305 a 132 c 177 g 256 t
ORIGIN	
Query Match	44.1%; Score 870; DB 2; Length 870;
Best Local Similarity	100.0%; Pred. No. 0;
Matches	870; Conservative 0; Mismatches 0; Indels 0; Gaps 0.
QY	364 atgacgagtgctaaagagatagaacccaataatcgagagatgtaaaaaagtacacaaaaaatt 423
DB	1 ATGGCAGGCGCTAAAGAGATAGAACCCAAAATCGGAGGTGAAGTACACAAAAAATT 60
QY	424 actaaagcgaatggaatggtgtcgtcctcgaaaatgctgaacgcgaagacgcatgtct 483
DB	61 ACTTAAGCGATGGAATGGTTGCTGCTCGAATAATCCGTAAACGACAGCATGTCT 120
QY	484 tcttcacgaccttattcagaaaacaatgaacgtatattagccaegtcccaagaacag 543
DB	121 TCTTCCACGCCCTTATTACAGAAACAATACGTACGTGATTAAGCCACGCTTCCAAAGCAAG 180
QY	544 attggttaaacatcccatcttttagtggatcggaagtaaaaaagtgggcattgtct 603
DB	181 ATTGGTTACAAACATCCCATTTTATAGGATCGGAGGTGAATAAAATGGGCATGATTGT 240
QY	604 gttgcacacgacgtcggtgctcttctgtgttgacttaaacgtgaactgtgtttaaactgtat 663
DB	241 GTGTCCACGATCGTGGTCTTTGTGGTGGCTTAACGTGAACCTGTTTAAACTGTATT 300
QY	664 aatgaatgaagaatggaagaagaaaagatgttcggtccaattgagttaatcggttct 723
DB	301 AATGAATGAAGAATGGAAGAAAAGATGTTCCGTTCAATTGAGTTTAATCGGTTCT 360
QY	724 aaatctataacttttccaactctttgggaattaaatttaacccaagaattaggtatt 783
DB	361 AAATCATATCAACTTTTCCAAATCTTTGGGAATTTAAATTTAAACCAAGATTAGGTATT 420
QY	784 gttgatattcccccgtgttagacgattaatgtgttcaactctatctatgtatgtctat 843
DB	421 GGTGATATCTCCCTCTGTGTAGCAGTTAATTTGTTCACTCAATCTATGATTGATGCTAT 480
QY	844 aaaaaaggggaagttagatgtgtgttattagtttataacaatttlaacacagatgctg 903
DB	481 AAAAAAGGGAATAGATGTGTGTTAGTTTATTAACAAATTTATTAAACAGATGTG 540
QY	904 caaaagccagttatggaataaataatccaattccacaggaattagataatgttattaggc 963
DB	541 CAAAGCCAGTATTGGAATAATTAATTCCTTACGAAATTAAGATAATGATGAATTAGGC 600

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 28, 2001, 15:36:06 ; Search time 2621.23 Seconds
(without alignments)
11636.675 Million cell updates/sec

Title: US-09-545-199c-3
Perfect score: 1972
Sequence: 1 agcgggcattgctcagf.....gtctattcttcgtgataa 1972

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 1344157 seqs, 773874588 residues

Word size : 25

Total number of hits satisfying chosen parameters: 12

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database :

GenEmbl:*
1: gb_ba1:*
2: gb_ba2:*
3: gb_ba3:*
4: gb_in1:*
5: gb_in2:*
6: gb_in3:*
7: gb_om:*
8: gb_ov:*
9: gb_pat1:*
10: gb_pat2:*
11: gb_ph:*
12: gb_pl1:*
13: gb_pl2:*
14: gb_pl3:*
15: gb_pl4:*
16: em_ba1:*
17: em_ba2:*
18: em_fun:*
19: em_htgo_hum:*
20: em_htgo_inv:*
21: em_htgo_rtd:*
22: em_htg_hum1:*
23: em_htg_hum2:*
24: em_htg_hum3:*
25: em_htg_hum4:*
26: em_htg_hum5:*
27: em_htg_hum6:*
28: em_htg_hum7:*
29: em_htg_hum8:*
30: em_htg_inv1:*
31: em_htg_inv2:*
32: em_htg_other:*
33: em_htg_rtd:*
34: em_hum1:*
35: em_hum2:*
36: em_hum3:*
37: em_hum4:*
38: em_hum5:*
39: em_hum6:*
40: em_hum7:*
41: em_in:*
42: em_om:*
43: em_or:*

44: em_ov:*
45: em_pat:*
46: em_ph:*
47: em_pl:*
48: em_rtd:*
49: em_sts:*
50: em_sy:*
51: em_un:*
52: em_vl:*
53: gb_sts1:*
54: gb_sts2:*
55: gb_sts3:*
56: gb_sy:*
57: gb_un:*
58: gb_vl1:*
59: gb_vl2:*
60: gb_hcg1:*
61: gb_hcg2:*
62: gb_hcg3:*
63: gb_hcg4:*
64: gb_hcg5:*
65: gb_hcg6:*
66: gb_hcg7:*
67: gb_hcg8:*
68: gb_hcg9:*
69: gb_hcg10:*
70: gb_hcg11:*
71: gb_hcg12:*
72: gb_hcg13:*
73: gb_hcg14:*
74: gb_hcg15:*
75: gb_hcg16:*
76: gb_hcg17:*
77: gb_hcg18:*
78: gb_hcg19:*
79: gb_hcg20:*
80: gb_hcg21:*
81: gb_hcg22:*
82: gb_hcg23:*
83: gb_hcg24:*
84: gb_hcg25:*
85: gb_pr1:*
86: gb_pr2:*
87: gb_pr3:*
88: gb_pr4:*
89: gb_pr5:*
90: gb_pr6:*
91: gb_pr7:*
92: gb_pr8:*
93: gb_pr9:*
94: gb_rtd:*
95: gb_rtd2:*
96: gb_in4:*
97: gb_pr10:*
98: em_ba3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1674	84.9	10229	1	AE006185 Pasteurel
2	870	44.1	870	2	AF237922 Pasteurel
3	44	2.2	1374	9	AR089428 Sequence
4	44	2.2	13828	3	U32730 Haemophilus
5	35	1.8	307	1	AF032869 Actinobac
6	32	1.6	793	1	AF037146 Salmonell
7	32	1.6	793	1	AF037147 Salmonell
8	32	1.6	793	1	AF037148 Salmonell

9 32 1.6 793 1 AF037149
10 32 1.6 793 1 AF037150
11 26 1.3 7307 2 BACATPA M20255 B.megaterium
c 12 25 1.3 34699 1 AF008210 AF008210 Buchnera

ALIGNMENTS

RESULT 1
AE006185 10229 bp DNA BCT 08-MAR-2001
LOCUS Pasteurella multocida PM70 section 152 of 204 of the complete
DEFINITION genome.
ACCESSION AE006185 AE004439
VERSION AE006185.1 GI:12721868
KEYWORDS
SOURCE Pasteurella multocida.
ORGANISM Pasteurella multocida
Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
Pasteurella.
REFERENCE 1 (bases 1 to 10229)
AUTHORS May,B.J., Zhang,Q., Li,L.L., Paustian,M.L., Whittam,T.S. and
Kaput,V.
TITLE Complete genomic sequence of Pasteurella multocida, PM70
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 98 (6), 3460-3465 (2001)
MEDLINE 21145866
REFERENCE 2 (bases 1 to 10229)
AUTHORS Zhang,Q. and Kaput,V.
TITLE Direct Submission
JOURNAL Submitted (24-OCT-2000) Department of Veterinary Pathobiology,
University of Minnesota, 1971 Commonwealth Ave., St. Paul, MN
55108, USA

FEATURES

source location/Qualifiers

1..10229 /organism="Pasteurella multocida"

/strain="PM70"

/db_xref="taxon:747"

275..2164

/gene="g1d4"

/note="PM1485"

275..2164

/gene="g1d4"

/codon_start=1

/transl_table=1

/product="g1d4"

/protein_id="AAK03569.1"

/db_xref="GI:12721869"

/translation="MFTYENYDVIVIGGHNATGALAPARMGLKTLILHNVDTLQ

MSCPVAIGIGIKGHLVREIDAMGIMATADQAGIOFTLNSSKGPVAVRATRADRV

LYRQAVIALENQENLIDFOEVDIILDQDRCVGVTKMGLPKHAKAVILTAGTFLS

GKHIGLENTGGRAGDPASVYMLADRLRELRLVDRLKTGPRRIDARTIDEVLAQ

HGDKLELVYFSPMGSDVQHPRIQCFITHTNDELRVNNLDRSPMAGITIEGIPRY

CPSTEDKVMRSEKNSHQITLDEPGILSNELTFNGISTSLPFDVQAKTVNSKMEKA

RIKPGVAIEVDYDPDLKPTLETKSIRGLFFAGQNGTGYEAGGGLAIDNNG

LFVQEKAMPFRPDQAYIGVLYDLCITGKEPRVTSRAEYRLRLRENDADSLRP

IAHQGLIDEKRAKRFNOKMENIELEORRLMLHPREYLDANKVIGSPVLRAS

GEDLRRPEMYOITLSTLTPROPAMDQEAQVEYIAIKQGYIEHOOEELAROKRH

STALPAHFDIVYSGLSNEVRAKLEQHRPYSIGASRLSGVTAPAIISILVLSKKQM

LKRG"

2164..2796

/gene="g1d4"

/note="PM1486"

2164..2796

/gene="g1d4"

/codon_start=1

/transl_table=1

/product="g1d4"

/protein_id="AAK03570.1"

/db_xref="GI:12721870"

/translation="MTNLEQQLSQKLELLKQTAISTDOQKQOLIKVILLNKRNA

YNLTSVDRPMELVKHILDSVYSPHLQGHFIDVGTGPGLPGLAVVNPKNHVL

DSLGRISFIRNAIREIGLDNVEAVLSRVEEYHPEQPPDGLSRASFSLKDMTDCOH

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

Oy 964 gaagaacaaagttcggatattatattacgaaccgtatgcgaagtattatagaatt 1023
 |||||||
 Db 601 GAAAGAAAAACAAGTTGGGATATTATATTTCGACACCGATGCGCAAGTATTATAGATAAT 660
 Oy 1024 ttattgttcgttaltttagaatctcaggtttlcaagcagcagttgaaacctgtctc 1083
 |||||||
 Db 661 TTATGTGTTCCGTTATTATTAGTAATCTCAGGTTTATCAAGCAGAGCTTGAAACCTTGCTTC 720
 Oy 1084 gagcaagccgctcgaatggttcgccatgaaagcagcaacagataacgcaggttaactt 1143
 |||||||
 Db 721 GAGCAAGCCGCTCGAATGCTCCGATGAAAGACAGAACGAATACGACAGTAACCTTAAT 780
 Oy 1144 aatggttacagttgactatatacaaaagctgtcgaacgaagtattacaataaataat 1203
 |||||||
 Db 781 AATGAGTTACAGATTAGTCTATATACAAAGCTCCTCANGCAAGATTTTACAAATGATTAAT 840
 Oy 1204 gaattgttcgcgcgtgcagcagcaatttaa 1233
 |||||||
 Db 841 GAAATTGTTGCCGCGTGCGACAGCACTATTAA 870
 RESULT 3
 LOCUS AR089428 1374 bp DNA PAT 07-SEP-2000
 DEFINITION Sequence 187 from patent US 5994066.
 ACCESSION AR089428
 VERSION AR089428.1 GI:10016185
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE
 AUTHORS 1 (bases 1 to 1374)
 TITLE Species-specific and universal DNA probes and amplification primers
 to rapidly detect and identify common bacterial pathogens and
 associated antibiotic resistance genes from clinical specimens for
 routine diagnosis in microbiology laboratories
 PATENT: US 5994066-A 187 30-NOV-1999;
 JOURNAL Location/Qualifiers
 FEATURES 1. 1374
 SOURCE /organism="Unknown"
 BASE COUNT 425 a 238 c 312 g 399 t
 ORIGIN
 Query Match 2.2% Score 44: DB 9; Length 1374;
 Best Local Similarity 100.0%: Pred. No. 1.Se-12;
 Matches 44: Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 1857 gttatggtcaatgaacgagccaccaggtacggttactggt 1900
 |||||||
 Db 607 GTTATGTCTCAATGACAGCAGCCACGAGTAACGTTTACGTGT 650
 RESULT 4
 LOCUS U32730 13828 bp DNA BCT 29-MAY-1998
 DEFINITION Haemophilus influenzae Rd section 45 of 163 of the complete genome.
 ACCESSION U32730.L2023
 VERSION U32730.1 GI:3212191
 KEYWORDS
 SOURCE Haemophilus influenzae Rd.
 ORGANISM Haemophilus influenzae Rd
 Bacteria: Proteobacteria; gamma subdivision: Pasteurellaceae:
 Haemophilus.
 1 (bases 1 to 13828)
 Fleischmann, R.D., Adams, M.D., White, O., Clayton, R.A.,
 Kirkness, E.F., Kerlavage, A.R., Bult, C.J., Tomb, J., Dougherty, B.A.,
 Merrick, J.M., McKenney, K., Sutton, G.G., Fitzhugh, W., Fields, C.A.,
 Gocayne, J.D., Scott, J.D., Shirley, R., Liu, L.I., Glodek, A.,
 Kelley, J.M., Weidman, J.F., Phillips, C.A., Spriggs, T., Hedblom, E.,
 Cotton, M.D., Utterback, T., Hanna, M.C., Nguyen, D.T., Saudek, D.M.,

Haemophilus influenza

REFERENCE	TITLE	REMARK	COMMENT	FEATURES
JOURNAL AUTHORS TITLE JOURNAL	INFLUENZAe Rd Science 269 (5223), 496-512 (1995)			
JOURNAL AUTHORS TITLE JOURNAL	2 (bases 1 to 13828) Tatsov R.L., Mushagian A.R., Bork P., Brown N.P., Hayes W.S., Borodovsky M., Rudd K.E. and Koonin E.V. Metabolism and evolution of Haemophilus influenzae deduced from a whole-genome comparison with Escherichia coli Curr. Biol. 6 (3), 279-291 (1996)			
JOURNAL AUTHORS TITLE JOURNAL	3 (bases 1 to 13828) White O., Clayton R.A., Kerlavage A.R. and Fleischmann R.D. Direct Submission Submitted (25-JUL-1995) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA			
JOURNAL AUTHORS TITLE JOURNAL	4 (bases 1 to 13828) White O., Clayton R.A., Kerlavage A.R. and Fleischmann R.D. Direct Submission Submitted (27-SEP-1997) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA			
JOURNAL AUTHORS TITLE JOURNAL	5 (bases 1 to 13828) White O., Clayton R.A., Kerlavage A.R., Fleischmann R.D., Peterson O., Hickey E., Dodson R. and Gwin M. Direct Submission Submitted (28-MAY-1998) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA			
JOURNAL AUTHORS TITLE JOURNAL	On Jun 12, 1998 this sequence version replaced gi:1573448. Location/Qualifiers			
JOURNAL AUTHORS TITLE JOURNAL	1.13828 /organism="Haemophilus influenzae Rd" /db_xref="taxon:11421" 65.1168 /gene="HI0470" 65.1168 /gene="HI0470" /note="similar to SP:P06986 GB:U02071 GB:X03416 PID:41695 PID:41710 percent identity: 60.40; identified by sequence similarity; putative" /codon_start=1 /transl_table=11 /product="histidinol-phosphate aminotransferase (hisc)" /protein_id="AAC22129.1" /db_xref="GI:1573449" transalation="MNDYSHIIQMKKTTITLTSRONIQAIPYOSAKRLGNGT IWNANETPSPEFOLSGKIDNRYRPOPOVAVYAGVSTENLVYFEGEGEEL IHTFCEKODALIFCPPTGYAVASRAGVSKYPLTDNDLMIPETKNIINLYK VIVVCSPNPTGKLLKQSDILDLQTTAGATVYVDEATIEFCEPSAVITLNNYHL AIIITRSKAPALGKCGFLANPELDILISKVIAPIPIVPSADIAEOLRPSNAT VOLTELLSRQMLAKALLVTHQVKEVSESEYLLIKQNGQAVFRAKLMWEGIIIR DOKETLHONCIIITVGTRECEKVEALEYEV" 1280.2368 /gene="HI0471" 1280.2368 /gene="HI0471" 1280.2368 /gene="HI0471" /note="similar to SP:P06987 GB:X03416 PID:41696 PID:41711 GB:U00096 percent identity: 64.97; identified by sequence similarity; putative" /codon_start=1 /transl_table=11 /product="imidazoleglycerol-phosphate dehydratase / histidinol-phosphatase (hisb)" /protein_id="AAC22130.1"			

```

/db_xref="GI:1573450"
/translation="MQPTLFIDRDGTLIDPKTDFOIDSLEKLEKLEPVYPLALLKKA
/translation="MQPTLFIDRDGTLIDPKTDFOIDSLEKLEKLEPVYPLALLKKA
KREIVYSMODGLTDAFQPTDFKPHNYMALFESOGITFDEVLICPHKPEENCICR
KPKILDLHKYIKRNLEEDIRSFVIGDRETDVLAELGIRAIQYDPOKMMDLIETKL
LGEYTNLDBHRRPRAEVDIRKTEIDKQVWLDEAGVEITEIGVFEDHMDLIATLH
GGERNNVQCKGDLWIDENHTVDEDTALALQALQKAVGDRGIRARGCVLPMDECKAE
ALDLGSRPIKPKFNACFKRDVGFSTSETEHFFQSLSFSLATLHLNVTGNNDHRIE
SLEKAFGRTRQAIRIEGNEPSSKGVLT"
/gene="H10472"
2434..3033
/gene="H10472"
2434..3033
/note="similar to SP:P10375 PID:41712 GB:U00096
PID:1736701 PID:1736710 percent identity: 55.90;
identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="amidotransferase (hisH)"
/protein_id="AAC22131.1"
/db_xref="GI:1573451"
/translation="MINITIIDPGCANLSSVKAFAFRLGYNTEITFDLNKIKSADKLI
LSCVGTANAMAYNLQEROLIEFTONTLOPVLICIGMOTERSEGNVPTLNLISGK
TRNIPDTGLPLPMGNRRVQFYKNCPLFGCIYONSHEFYHSAVSPNHSVAISVYG
VNSAIAIAKENFYGVQFHPERSGKNGALLLNKFEKVPF"
3069..3818
/gene="H10473"
3069..3818
/note="similar to SP:P10371 PID:41713 GB:U00096
PID:1736702 PID:1736711 percent identity: 62.86;
identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="phosphoribosylformimino-5-aminoimidazole
carboxamide ribotide isomerase (hisA)"
/protein_id="AAC22132.1"
/db_xref="GI:1573452"
/translation="MKOSIITIPALDLINGOVVRLHOGDYAKQTTYSNDPIKQFNDYVR
OGKAKHLVDLTCAGKNSORRQALIGKIVATCGKQVGGIGRTEDVDADLLAVGNR
VVGSTVAHRSVKNMFIKYGAKEVYALDVNINSGKIVALSIGMOESGVLLETL
IEPQVYGLQVYCTDISRGTGSGNIGXOPEICERYPPIQSSGIGSLADIEAL
KGTGSGVIVGRALBEGFTLSERAIKQWNG"
3800..4576
/gene="H10474"
3800..4576
/note="similar to SP:P10373 PID:41714 PID:598474 GB:U00096
PID:1736704 percent identity: 82.03; identified by
sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="hisF cyclase (hisF)"
/protein_id="AAC22133.1"
/db_xref="GI:1573453"
/translation="MLAKRIIPCLIDVRDGVKGVQFRNHETICTDYLPLAQRVQESA
DELIVFYITASSDGRKSVWERIAVVIDIPFCVAGICTTEAEKLFARAGAKISL
NSPLADPTLISLADRFQVQALVAGIDSEFTGVMWQYDGTGRTKROTHWOLL
DMWEYQCGAGETIVLMMNODGLRNGYDIAOLKLVGVCVRPLIAGAGAKWVHFRD
AETAKYDGLASVFFHKQILEIGELKSYLVQSIERSE"
4576..5241
/gene="H10475"
4576..5241
/note="similar to SP:P06989 GB:D43637 GB:U02072 GB:X03974
PID:41700 percent identity: 61.22; identified by sequence
similarity; putative"
/codon_start=1
/transl_table=11
/product="phosphoribosyl-AMP cyclohydrolase /
phosphoribosyl-ATP pyrophosphohydrolase (hisIb)"
/protein_id="AAC22134.1"
/db_xref="GI:1573454"
/translation="MNITKIDMKVNGLLPVYIIONISTREVLMGYNNEALTTTKE

```

```

RKVPFSRTRKRLMTKGEISGNFLNVEEMSLDCCNDTLLILVDPIGATCHGEVSCFH
DETSPOSENKQOAFNMWAMFIKEQHLKEKKADPSNSYTAIHLAKGTRKIAQKQEE
GVETALAAVADQKAAVISEATDVLVHLVTVLHNDQLQWVEITAKIQEHRHOGIGLHPEG
GNK"
5323..5475
/gene="H10476"
5323..5475
/note="H10476"
5323..5475
/note="hypothetical protein; identified by Genemark;
putative"
/codon_start=1
/transl_table=11
/product="H. influenzae predicted coding region H10476"
/protein_id="AAC22144.1"
/db_xref="GI:1573464"
/translation="MVIKCIDQKQNLNIIIFLLKQXSKEDSKKFTYKRYLQTVN
VYIOLS"
5505..6707
/gene="H10477"
5505..6707
/note="H10477"
5505..6707
/note="similar to GB:M23240 SP:P18199 PID:148089 GB:U00096
PID:1736569 percent identity: 47.24; identified by
sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="tyrosine-specific transport protein (tyrP)"
/protein_id="AAC22135.1"
/db_xref="GI:1573455"
/translation="NKKTVGSTLLVAGTMIGAGMLAMPLTSAGIGFTVLVLLGLMA
LTFPSALVEVEICOTAESDAGIGTAEQYFGKGRIRIAVAVLIFPLAAVAVISGG
SLIKDLIPESFGKVSLLFTVYFGSEIYGFHSVDKINRVLEFVMAAFVAVLSML
PEIKFNLNATPIPKALILISASPVFTFAGFHSIPSLNKYLDGNVKALRESVLVSA
ITLCAYILWOLSTHGLTQNEFLQILKEKRTNLGLYKATPATGNSVAVISAVLFTSL
ALITSPVGVGLGLEICIEDLKRSEFVNTAGRTSLGLTTIPLPVPLFTPREGTLALG
YAGOMPFAFVAVLPVSLVWKARRAHANLPLKVGWMLTILVYLVLITSLTSPFAIRA
GYLPFVYG"
complement(6893..7321)

gene
Query Match
Best Local Similarity 100.0%; Prev. No. 1,6e-12;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1857 gttatggtcaaatgaacgagccacggtacggtttacgtt 1900
|||||
Db 8118 GTTATGTCCTAATGAACGACGACGATACCGTTACGTGT 8075

RESULT 5
AF032869 307 bp DNA BCT 27-NOV-1997
LOCUS Actinobacillus pleuropneumoniae ATP synthase gene, partial cds.
DEFINITION AF032869
ACCESSION AF032869
VERSION AF032869.1 GI:2641642
KEYWORDS
SOURCE
ORGANISM
Actinobacillus pleuropneumoniae.
Actinobacillus pleuropneumoniae
Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
Actinobacillus.
1 (bases 1 to 307)
REFERENCE
AUTHORS Baeckeister,C.X. and Fenwick,B.W.
TITLE Direct Submission
JOURNAL Submitted (03-NOV-1997) Diagnostic Medicine/Pathobiology, Kansas
State University, 1800 Denison Ave, Manhattan, KS 66506-5605, USA
FEATURES
source
1..307
/organism="Actinobacillus pleuropneumoniae"
/strain="4074"
/serotype="1"
/db_xref="taxon:715"
1..>307
/codon_start=1
/transl_table=11

```

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 28, 2001, 14:58:41 ; Search time 74.92 Seconds
(without alignments)
4962.936 Million cell updates/sec

Title: US-09-545-199C-3

Perfect score: 1972

Sequence: 1 agcggcattgctgcacg.....gtctattctctgtgataa 1972

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 324599 seqs, 9465562 residues

Total number of hits satisfying chosen parameters: 649198

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents, NA:*

- 1: /cgn2.6/ptodata/1/ina/5A.COMB.seq:*
- 2: /cgn2.6/ptodata/1/ina/5B.COMB.seq:*
- 3: /cgn2.6/ptodata/1/ina/6A.COMB.seq:*
- 4: /cgn2.6/ptodata/1/ina/6B.COMB.seq:*
- 5: /cgn2.6/ptodata/1/ina/PCURS.COMB.seq:*
- 6: /cgn2.6/ptodata/1/ina/backfiles1.seq:*

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	555.6	28.2	1374	2	US-08-743-637B-187
2	48.8	2.5	7218	1	US-08-232-463-14
3	41.8	2.1	872	4	US-08-998-416-487
4	40.4	2.0	3095	6	5231168-1
5	40.2	2.0	2107	4	US-08-235-836C-73
6	38.8	2.0	1899	1	US-08-634-331-2
7	38.6	2.0	1991	4	US-08-235-836C-69
8	38.6	2.0	1991	4	US-08-235-836C-77
9	38	1.9	5361	4	US-08-973-462-2
10	38	1.9	6152	4	US-08-973-462-1
11	36.4	1.8	1689	1	US-07-991-867B-41
12	36.4	1.8	1689	2	US-08-544-332-41
13	36.4	1.8	8457	1	US-07-991-867B-1
14	36.4	1.8	8457	2	US-08-544-332-1
15	36.2	1.8	9636	1	US-08-232-170B-1
16	36.2	1.8	10395	1	US-08-245-809-3
17	36.2	1.8	10396	1	US-08-245-809-5
18	36.2	1.8	10798	1	US-08-107-748-2
19	36.2	1.8	10798	5	PCT-US92-01385-2
20	36.2	1.8	10965	1	US-08-107-748-4
21	36.2	1.8	10965	5	PCT-US92-01385-4
22	36	1.8	1152	2	US-08-286-819A-13
23	36	1.8	1152	3	US-08-980-357-13
24	36	1.8	2081	4	US-08-235-836C-71
25	36	1.8	7225	4	US-08-286-819A-15
26	36	1.8	7225	3	US-08-980-357-15
27	36	1.8	10851	2	US-08-286-819A-16

28	36	1.8	10851	3	US-08-980-357-16	Sequence 16, Appl
29	36	1.8	19124	2	US-08-487-826B-13	Sequence 13, Appl
30	35.6	1.8	959	2	US-08-634-331-1	Sequence 1, Appl
31	35.6	1.8	959	6	5508165-1	Patent No. 5508165
32	35.6	1.8	2570	2	US-09-056-075-2	Sequence 2, Appl
33	35.4	1.8	1146	4	US-09-277-716-21	Sequence 21, Appl
34	35.4	1.8	2319	2	US-08-966-389-5	Sequence 5, Appl
35	35.4	1.8	2319	2	US-09-103-509-5	Sequence 5, Appl
36	35.4	1.8	2319	2	US-09-102-644-5	Sequence 5, Appl
37	35.4	1.8	2319	2	US-09-218-032-5	Sequence 5, Appl
38	35.4	1.8	3345	2	US-08-966-389-8	Sequence 8, Appl
39	35.4	1.8	3345	2	US-09-103-509-8	Sequence 8, Appl
40	35.4	1.8	3345	2	US-09-102-644-8	Sequence 8, Appl
41	35.4	1.8	3345	2	US-09-218-032-8	Sequence 8, Appl
42	35.4	1.8	19124	2	US-08-487-826B-13	Sequence 13, Appl
43	34.8	1.8	1035	2	US-08-417-210A-111	Sequence 111, App
44	34.8	1.8	2542	4	US-09-330-317B-9	Sequence 9, Appl
45	34.8	1.8	7101	1	US-08-480-604A-9	Sequence 9, Appl

ALIGNMENTS

RESULT 1
US-08-743-637B-187
Sequence 187, Application US/08743637B
Patent No. 5994066
GENERAL INFORMATION:
APPLICANT: BERGERON, Michel G.
APPLICANT: PICARD, Francois J.
APPLICANT: COELLETTTE, Marc
APPLICANT: ROY, Paul H.
TITLE OF INVENTION: SPECIES-SPECIFIC AND UNIVERSAL DNA
TITLE OF INVENTION: PROBES AND AMPLIFICATION PRIMERS TO RAPIDLY DETECT AND
TITLE OF INVENTION: IDENTIFY COMMON BACTERIAL PATHOGENS AND ASSOCIATED
TITLE OF INVENTION: ANTIBIOTIC RESISTANCE GENES FROM CLINICAL SPECIMENS
NUMBER OF SEQUENCES: 273
CORRESPONDENCE ADDRESSES:
ADDRESSEE: QUARLES & BRADY
STREET: 411 EAST WISCONSIN AVENUE
CITY: MILWAUKEE
STATE: WISCONSIN
COUNTRY: USA
ZIP: 53202-4497
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/743,637B
FILING DATE: 04-NOV-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/526,840
FILING DATE: 11-SEP-1995
ATTORNEY/AGENT INFORMATION:
NAME: BAKER, Jean C.
REGISTRATION NUMBER: 35,433
REFERENCE/DOCKET NUMBER: 850586.90012
TELEPHONE: (414) 277-5000
TELEFAX: (414) 277-5591
INFORMATION FOR SEQ ID NO: 187:
SEQUENCE CHARACTERISTICS:
LENGTH: 1374 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: Haemophilus Influenzae
US-08-743-637B-187

[illegible]

STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,463
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/935,313
FILING DATE:
APPLICATION NUMBER: EP 91 114 300.6
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/114 IMMU
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300
TELEFAX: (703)683-4109
TELEX: 899143
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 7218 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: PTZgpt-fls
US-08-232-463-14

Query Match	2.5%	Score 48.8	DB 1	Length 7218
Best Local Similarity	5.7%	Pred. No. 0.0029		
Matches	23	Conservative 212	Mismatches 169	Indels 0
				Gaps 0
Qy	289	gfcattaaagcatttgtagtgcctcaagaacacgctgctglaagtltaacactt	348	
Db	1483	GPATTAACCTATCTTATGCAAGTAGTTAAAGACATAGACAGATTGTGACRRRRRRRR	1424	
Qy	349	aaatgagagacacaatgycaggtgctaaagagataagaacacaatcgcgagtgtaaa	408	
Db	1423	RR	1364	
Qy	409	agtacacaaaatctctaagcgcgtggaatgctgtgctccctcgaaaagcgtlaaacg	468	
Db	1363	RR	1304	
Qy	469	caagaagcgtgctcttcacgcccattacagaacaatacgtlaacgtgattagcac	528	
Db	1303	RR	1244	
Qy	529	gttccaaagcaacgatlgtgltacaagatccattttagtgcgcgaagtaaaaaa	588	
Db	1243	RR	1184	
Qy	589	gtggcgatgattgtgtgtccacagatcgtgccttggtggtggtctaagcgtbaactg	648	
Db	1183	RR	1124	
Qy	649	tttaaaactgtlataatgaaatgaaagaatggaagaaaaaaga	692	
Db	1123	RR	1080	

RESULT 3

US-08-998-416-487

; Sequence 487, Application US/08998416

; Patent No. 6239264

```

: GENERAL INFORMATION:
: APPLICANT:  Philippson, Peter
: APPLICANT:  Pohlmann, Rainer
: APPLICANT:  Steiner, Sabine
: APPLICANT:  Mohr, Christine
: APPLICANT:  Wendland, Jürgen
: APPLICANT:  Knechtle, Philipp
: APPLICANT:  Reibschung, Corinne
: TITLE OF INVENTION:  GENOMIC DNA SEQUENCES OF ASHBVA GOSSIPIT
: TITLE OF INVENTION:  AND USES THEREOF
: NUMBER OF SEQUENCES:  1152
: CORRESPONDENCE ADDRESS:
: ADDRESSEE:  No. 6239264artis Corporation
: STREET:     3054 Cornwallis Road
: CITY:       Research Triangle Park
: STATE:      No. 6239264th Carolina
: COUNTRY:    USA
: ZIP:        27709
:
: COMPUTER READABLE FORM:
: MEDIUM TYPE:  Floppy disk
: COMPUTER:     IBM PC compatible
: OPERATING SYSTEM:  PC-DOS/MS-DOS
: SOFTWARE:     Patentin Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER:  US/08/998,416
: FILING DATE:       24-DEC-1997
: CLASSIFICATION:    435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER:  CH 0016/97
: FILING DATE:       31-DEC-1996
: ATTORNEY/AGENT INFORMATION:
: NAME:            Meigs, J. Timothy
: REGISTRATION NUMBER:  38,241
: REFERENCE/DOCKET NUMBER:  PF/5-30306/A/CGC1976
: TELECOMMUNICATION INFORMATION:
: TELEPHONE:       919-541-8587
: TELEFAX:        919-541-8689
: INFORMATION FOR SEQ ID NO:  487:
: SEQUENCE CHARACTERISTICS:
: LENGTH:        872 base pairs
: TYPE:           nucleic acid
: STRANDEDNESS:  single
: TOPOLOGY:      linear
: MOLECULE TYPE:  DNA (genomic)
: ORIGINAL SOURCE:
: ORGANISM:       PAG1344UP
:
: US-08-998-416-487
:
: Query Match          2.1%; Score 41.8; DB 4; Length 872;
: Best Local Similarity 46.9%; Pred. No.0.079;
: Matches 161; Conservative 0; Mismatches 179; Indels 3; Gaps 1.

```

```

      YQ      948 taatgatgaattagcgaagaagaacaacgattgggcattataatt 990
          ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
      Db      666 TAATTAATAAGAAATGGAAGANACACTCTCAANAATTAAATT 728

RESULT      4
5231168-1
Patent No. 5231168
Applicant: DIEZIEL, MORTEN; BORRE, MARTIN; JEPSEN, SOREN;
VUUST, JENS; RIENECK, KLAUS; WIND, ANNETTE; JAKOBSEN, PALLE H.;
TITLE OF INVENTION: MALARIA ANTIGEN
NUMBER OF SEQUENCES: 19
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/409,658
FILING DATE: 18-SEP-1989
SEQ ID NO: 1:
LENGTH: 3095
5231168-1

Query Match      2.0%; Score 40.4; DB 6; Length 3095;
Best Local Similarity 49.3%; Pred. No. 0.31;
Matches 133; Conservative 0; Mismatches 136; Indels 1; Gaps 1;

OY      1129 gcaagttaactaatthaagggttacagtttagtctctaacaagctcgccgaagcatgtt 1188
      Db      115 gaagaataaattagtttgataaagatgaatgaaggtgttttgagaagccaccatgaataattt 174
OY      1189 acaaatgaattaataatgaataattggttcggtgcagcagcaattttaacaataagagatcg- 1247
      Db      175 tcatctgaagtaagaagtaaacctcgaattaaatgaataattgttgatctgcacaaaagt 234
OY      1248 gtaatgccaactgtgaaaaaatttgtlaacaatcaltcggtgcggttatattgacgttgaattccca 1307
      Db      235 gtaactgaacctgctgacaatgaagaagttgtatctgaagaagaagcaaccctgaaccagct 294
OY      1308 caagtgagcagraccaaaagtatatgatgccttaaatgltgnaaacaggttgtaattgaa 1367
      Db      295 gaaatgaagaagaagtagtataagaagaagctcatcgcgaagaagaattgtactgnaacaaat 354
OY      1368 gtacacacaattagtggtgcgttgaatt 1397
      Db      355 gatgaagaatcagtgtaagaagtcgattagtt 384

RESULT      5
US-08-235-836C-73
Sequence 73, Application US/08235836C
Patent No. 6248562
GENERAL INFORMATION:
APPLICANT: Dunn, John J.
APPLICANT: Luft, Benjamin J.
TITLE OF INVENTION: No. 6248562el Chimeric Proteins Comprising
TITLE OF INVENTION: Borrelia Polypeptides and Uses Therefor
NUMBER OF SEQUENCES: 144
CORRESPONDENCE ADDRESS:
ADDRESSSEE: Brookhaven National Laboratory
STREET:
CITY: Upton
STATE: NY
COUNTRY: USA
ZIP: 11973
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/235,836C
FILING DATE: 29-APR-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/148,191

```


US-08-235-836C-77

Query Match	2.0%	Score	38.6	DB	4	Length	1991
Best Local Similarity	48.1%	Pred. No.	0.76				
Matches	139	Conservative	0	Mismatches	149	Indels	1
						Gaps	1

OY	1110	gaagcgcaacagatatacgagtaacttaataaggtacagttagttcattaca	116
Db	654	GAATGATTCAGGTTTACTTTGCCAGACGATTATTACGACATTTCAAGGCCAAATCTCATTA	713
OY	1170	agctgcgaagcaagtaattacaatgaattaaatgaattgttcgcggtgcagcaagcaat	1221
Db	714	AGCGATGACAGATTAATTAATTGATATTGTAATTAGATAATTTTCATGAAAGTGATTTCCAAATAT	773
OY	1230	ttacaatatgagatctcggttaatggcgcaactcggaaaaattgtacaatatcatcggtgc	128
Db	774	AACGGAACATATTGACGATTTAAAGGGATCGCTTGGAAAAGTACGATGACAGTAAACAGCATTA	833
OY	1289	tattgacgtgaattcccaacaagatcagatccacaagaatataatgattcccttaaatgtga	134
Db	834	AAAAGACATTGAAAGTCAGCTTGCTTAAGAAAGAAACAAAGAAAGAAATTAAGATTAATAA	893
OY	1349	aacaggttcgtactcgtgaattcaacaacaatgaatgtgtgtgtgtgtgtc	1397
Db	894	GGCATTGATCTTGATTAACCTCAACCAAAATTAATGATTTTGCTGGAAGAT	942

RESULT 9
US-08-973-462-2
; Sequence 2, Application US/08973462B

```

1  GENERAL INFORMATION:
2  APPLICANT: DRUILHE, PIERRE
3  APPLICANT: DAUBERSTES, PIERRE
4  TITLE OF INVENTION: MALARIAL PRE-ERYTHROCYTIC STAGE POLYPEPTIDE MOLECULES
5  FILE REFERENCE: 0660-0125-0 PCT
6  CURRENT APPLICATION NUMBER: US/08/973,462B
7  CURRENT FILING DATE: 1998-02-06
8  EARLIER APPLICATION NUMBER: PCT/FR96/00894
9  EARLIER FILING DATE: 1996-06-12
10 EARLIER APPLICATION NUMBER: FR 95/07007
11 EARLIER FILING DATE: 1995-06-13
12 NUMBER OF SEQ ID NOS: 29
13 SOFTWARE: PatentIn Ver. 2.0

```

```

; SOFTWARE: PatentIn Ver.
; SEQ ID NO 2
; LENGTH: 5361
; TYPE: DNA
; ORGANISM: P. falciparum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(5361)
; OS-08-973-462-2

```

Query Match	1.9%;	Score 38;	DB 4;	Length 5361;
Best Local Similarity	44.9%;	Pred. NO. 1.6;		
Matches 266; Conservative	0;	Mismatches 320;	Indels 6;	Gaps 3

OY 808 ttaattggttcagccaattctatgatgttgatcctaataaaagg-ggaagtatgatttgc 866
||| ||||| ||||| ||||| ||||| |||||
Db 3639 tgaaagtgttggaagaagcttttagaatgtgaaaaatatggaagaaggaattcaatgaaagatgc 3699

OY	867	gtaattcgtttatacaacaattttaaaccgaatgctgcgaagaagccgatcttggaanaatt	926
		" "	
Db	3699	tgttgaataataatgacatacttagccaacttatcttgaaagaactcaagatgaataagt	3758
OY	927	aattccattcacccaagattagataatgatatagcttagcgcaagaanaacaagtttggatta	986
		" "	
Db	3759	agaagcagattataataaagatatgtaaaaatttaagaagatttagaanaagcatattcaga	3818
OY	987	tatttcgaacctgatacgaaagtattattatagataattatttgttctgattattagatc	1046
		" "	
Db	3819	aggattctcaagaataataatagatgcgaagaatgatacatatgnaaaaaagttattgaagga	3878
OY	1047	tcaggtttaccaagcagtcagttgnaaaccttgctctgcgcaagccgctcgaaatgctgc	1106
		" "	
Db	3879	acatgatatatacgaacgcgcgttgtagtgtaagtgtgtgaattaaagatgtlcgaaagacaa	3938
OY	1107	cataaagcgcgaacacagataacgcgaagtgtaacttaattaatgattacagttagt---	1153
		" "	
Db	3939	gatcgaaaagaagtatctgattttaaagctcttgaaagaatataattaaagaagaanaaga	3998
OY	1164	taacaaagcttcgttaagcaagttattacaagaatttaaatgtaattgttgcgcgtgcagc	1223
		" "	
Db	3999	aatcaaaagaactctyaaagltgaatttagaagatataaagaattaaactaattgaaac	4058
OY	1224	agcaatttaacaagtagagatctgtaatagtgcacaactgynaanaattgltcaaacatcgt	1283
		" "	
Db	4059	agatattttgaagagaga--aaaagaattagaaaaagctcattgnaaaattcgaaaga	4116
OY	1284	gcggttattgscgttgaattcccacaagatgcagttaccnaaagtataatgagccctaatt	1343
		" "	
Db	4117	gaagctggaagaataaaagaactcttgaagcagatatatttaaaagaagatcttcattagaa	4176
OY	1344	gttaaaacaggtttgacttggaagttcaacaacacattaggtgtygtgtytag	1395
		" "	
Db	4177	gtcgaagaagaanaaaaaattagaagaagtatacgaattaaagaagaaggttag	4228
RESULT 10			
US-08-973-462-1			
Sequence 1, Application US/08973462B			
Patent No. 6191270			
GENERAL INFORMATION:			
APPLICANT: DRULIHE, PIERRE			
TITLE OF INVENTION: MARIARIAL PRE-ERTHROCYTIC STAGE POLYPEPTIDE MOLECULES			
FILE REFERENCE: 0660-0125-0 PCT			
CURRENT APPLICATION NUMBER: US/08/973.462B			
EARLIER FILING DATE: 1998-02-06			
EARLIER APPLICATION NUMBER: PCT/FR96/00894			
EARLIER FILING DATE: 1996-06-12			
EARLIER APPLICATION NUMBER: FR 95/07007			
NUMBER OF SEQ ID NOS: 29			
SOFTWARE: PatentIn Ver. 2.0			
SEQ ID NO 1			
LENGTH: 6152			
TYPE: DNA			
ORGANISM: P. falciparum			
US-08-973-462-1			

Query Match	Similarity	1.9%	Score 38;	DB 4;	Length 6152;
Best Local	Similarity	44.9%	Pred No. 1.77;		
Matches	266;	Conservative	0;	Mismatches 320;	Indels 6; Gaps 3
Qy	808	ttaattggttcagtcacattcatatgatgtgcttataaaagg-ggaagtatgtttc	866		
Db	3883	tgaaggtgttggaaggttttgaattgtaaaaataatcgaagaggaattatgaagaatc	3942		
Qy	867	gattttagttatacaaatcttatacaagatgtgcgaagcccgatttggaaaatt	926		
Db	3943	tattgaataatgaatcattactgcgaacttattgaagaatccagaggttlaattgaat	4002		

[illegible]

RESULT 11
 US-07-991-867B-41/C
 Sequence 41, Application US/07991867B
 Patent No. 5476781
 GENERAL INFORMATION:
 APPLICANT: Moyer, Richard W.
 APPLICANT: Hall, Richard L.
 APPLICANT: Grudl, Michael E.
 TITLE OF INVENTION: No. 5476781el Entomopoxvirus Expression System
 NUMBER OF SEQUENCES: 66
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: David R. Saliwanchik
 STREET: 2421 N.W. 41st Street, Suite A-1
 CITY: Gainesville
 STATE: FL
 COUNTRY: USA
 ZIP: 32606
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentln Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/07/991,867B
 FILING DATE: 12-DEC-1992
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: WO 92/14818
 FILING DATE: 12-FEB-1992
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/827,685
 FILING DATE: 30-JAN-1992
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/657,584
 FILING DATE: 19-FEB-1991
 ATTORNEY/AGENT INFORMATION:
 NAME: Saliwanchik, David R.
 REGISTRATION NUMBER: 31,794
 REFERENCE/DOCKET NUMBER: UFI14.C3

Search completed: October 28, 2001, 15:36:00
Job time: 2239 sec

```

RESULT 15
US-08-323-1708-1
: Sequence 1, Application US/083231708
: Patent No. 5733772
GENERAL INFORMATION:
APPLICANT: Williamson, Kim C.
APPLICANT: Kaslow, David C.
TITLE OF INVENTION: Cloning and Expression of Plasmidium
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/323,170B
FILING DATE: 13-OCT-1994
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/010,409
FILING DATE: 29-JAN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Quine, Jonathan A.
REGISTRATION NUMBER: P-41,261
REFERENCE/DOCKET NUMBER: 015280-113100US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 9636 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 149..9556
US-08-323-1708-1

```

QY	918	ggaataatcaattccattaccgaattagatatcatgatgaattaggcgaagaagaacaagt	977
Db	1003	GGAAGAAGAAAGCAGGAGAGGAGAAACAGGAAAGAGAAAGAGCAGCAGCAAGATTA	10652
QY	978	ttaggatataattaacgaancccgatgcgaagaatattatagataaatttatttggttcgtta	1037
Db	1063	TGATGATTATGGTTTAAAGGAAAAGTGCGGATATAAACAGCAAGAACCATTTACAAGAGAACCA	11222

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 28, 2001, 14:58:41 ; Search time 1501.77 Seconds
(without alignments)
12412.704 Million cell updates/sec

Title: US-09-545-199C-3
Perfect score: 1972
Sequence: 1 agcggccattgctcagf.....gtctattcttcgttgataa 1972

Scoring table: IDENTITY_NDC
Gapop 10.0 , Gapext 1.0

Searched: 10228115 segs, 4726426750 residues
Total number of hits satisfying chosen parameters: 20456230

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

```
EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_est4:*
5: gb_est5:*
6: gb_est6:*
7: gb_est7:*
8: gb_est8:*
9: gb_est9:*
10: gb_est10:*
11: gb_est11:*
12: gb_est12:*
13: gb_est13:*
14: gb_est14:*
15: gb_est15:*
16: gb_est16:*
17: gb_est17:*
18: gb_est18:*
19: gb_est19:*
20: gb_est20:*
21: gb_est21:*
22: gb_est22:*
23: gb_est23:*
24: gb_est24:*
25: gb_est25:*
26: gb_est26:*
27: gb_est27:*
28: gb_est28:*
29: gb_est29:*
30: gb_est30:*
31: gb_est31:*
32: gb_est32:*
33: gb_est33:*
34: gb_est34:*
35: gb_est35:*
36: gb_est36:*
37: gb_est37:*
38: gb_est38:*
39: gb_est39:*
40: gb_est40:*
41: gb_est41:*
42: gb_est42:*
43: gb_est43:*
44: em_esthum10:*
45: em_esthum11:*
46: em_esthum12:*
47: em_esthum13:*
48: em_esthum14:*
49: em_esthum15:*
50: em_esthum16:*
51: em_esthum17:*
52: em_esthum18:*
53: em_esthum19:*
54: em_esthum20:*
55: em_esthum21:*
56: em_esthum22:*
57: em_esthum23:*
58: em_esthum24:*
59: em_esthum25:*
60: em_esthum26:*
61: em_esthum27:*
62: em_esthum28:*
63: em_esthum29:*
64: em_esthum30:*
65: em_esthum31:*
66: em_esthum32:*
67: em_esthum33:*
68: em_esthum34:*
69: em_esthum35:*
70: em_esthum36:*
71: em_esthum37:*
72: em_esthum38:*
73: em_esthum39:*
74: em_esthum40:*
75: em_esthum41:*
76: em_esthum42:*
77: em_esthum43:*
78: em_esthum44:*
79: em_esthum45:*
80: em_esthum46:*
81: em_esthum47:*
82: em_esthum48:*
83: em_esthum49:*
84: em_esthum50:*
85: em_esthum51:*
86: em_esthum52:*
87: em_esthum53:*
88: em_esthum54:*
89: em_esthum55:*
90: em_esthum56:*
91: em_esthum57:*
92: em_esthum58:*
93: em_esthum59:*
94: em_esthum60:*
95: em_esthum61:*
96: em_esthum62:*
97: em_esthum63:*
98: em_esthum64:*
99: em_esthum65:*
100: em_esthum66:*
101: em_esthum67:*
102: em_esthum68:*
103: em_esthum69:*
104: em_esthum70:*
105: em_esthum71:*
106: em_esthum72:*
107: em_esthum73:*
108: em_esthum74:*
109: em_esthum75:*
110: em_esthum76:*
111: em_esthum77:*
112: em_esthum78:*
113: em_esthum79:*
114: em_esthum80:*
115: em_esthum81:*
116: em_esthum82:*
```

117: gb_est48:*
118: gb_est49:*
119: gb_est50:*
120: gb_est51:*
121: gb_est52:*
122: gb_est53:*
123: gb_est54:*
124: gb_est55:*
125: gb_est56:*
126: gb_est57:*
127: gb_est58:*
128: gb_est59:*
129: gb_est60:*
130: gb_est61:*
131: gb_est62:*
132: gb_est63:*
133: gb_est64:*
134: gb_est65:*
135: gb_est66:*
136: gb_est67:*
137: gb_est68:*
138: gb_est69:*
139: gb_est70:*
140: gb_est71:*
141: gb_est72:*
142: gb_est73:*
143: gb_est74:*
144: gb_est75:*
145: gb_est76:*
146: gb_est77:*
147: gb_est78:*
148: gb_est79:*
149: gb_est80:*
150: gb_est81:*
151: gb_est82:*
152: gb_est83:*
153: gb_est84:*
154: gb_est85:*
155: gb_est86:*
156: gb_est87:*
157: gb_est88:*
158: gb_est89:*
159: gb_est90:*
160: gb_est91:*
161: gb_est92:*
162: gb_est93:*
163: gb_est94:*
164: gb_est95:*
165: gb_est96:*
166: gb_est97:*
167: gb_est98:*
168: gb_est99:*
169: gb_est100:*
170: gb_est101:*
171: gb_est102:*
172: gb_est103:*
173: gb_est104:*
174: gb_est105:*
175: gb_est106:*
176: gb_est107:*
177: gb_est108:*
178: gb_est109:*
179: gb_est110:*
180: gb_est111:*
181: gb_est112:*
182: gb_est113:*
183: gb_est114:*
184: gb_est115:*
185: gb_est116:*
186: gb_est117:*
187: gb_est118:*
188: gb_est119:*
189: gb_est120:*

190: gb_est121:*
191: gb_est122:*
192: gb_est123:*
193: gb_est124:*
194: gb_est125:*
195: gb_est126:*
196: gb_est127:*
197: gb_est128:*
198: gb_est129:*
199: gb_est130:*
200: gb_est131:*
201: gb_est132:*
202: gb_est133:*
203: gb_est134:*
204: gb_est135:*
205: gb_est136:*
206: gb_est137:*
207: gb_est138:*
208: gb_est139:*
209: gb_est140:*
210: gb_est141:*
211: gb_est142:*
212: gb_est143:*
213: gb_est144:*
214: gb_est145:*
215: gb_est146:*
216: gb_est147:*
217: gb_est148:*
218: gb_est149:*
219: gb_est150:*
220: gb_est151:*
221: gb_est152:*
222: gb_est153:*
223: gb_est154:*
224: gb_est155:*
225: gb_est156:*
226: gb_est157:*
227: gb_est158:*
228: gb_est159:*
229: gb_est160:*
230: gb_est161:*
231: gb_est162:*
232: gb_est163:*
233: gb_est164:*
234: gb_est165:*
235: gb_est166:*
236: gb_est167:*
237: gb_est168:*
238: gb_est169:*
239: gb_est170:*
240: gb_est171:*
241: gb_est172:*
242: gb_est173:*
243: gb_est174:*
244: gb_est175:*
245: gb_est176:*
246: gb_est177:*
247: gb_est178:*
248: gb_est179:*
249: gb_est180:*
250: gb_est181:*
251: gb_est182:*
252: gb_est183:*
253: gb_est184:*
254: gb_est185:*
255: gb_est186:*
256: gb_est187:*
257: gb_est188:*
258: gb_est189:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.


```

RESULT 2
LOCUS   AUI75106      334 bp      mRNA      EST      21-MAR-2001
DEFINITION AUI75106 Rice panicle at ripening stage Oryza sativa cDNA clone
            E10189, mRNA sequence.
ACCESSION AUI75106
VERSION   AUI75106.1  GI:13421016
KEYWORDS EST.
SOURCE    Oryza sativa.
ORGANISM  Oryza sativa.
           Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
           Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
           Ehrhartoideae; Oryzaceae; Oryza.
REFERENCE 1 (bases 1 to 334)
AUTHORS   Sasaki,T. and Yamamoto,K.
TITLE     Rice cDNA from panicle at ripening stage (2001)
JOURNAL   Unpublished (2001)
COMMENT   Contact: Takuji Sasaki
           National Institute of Agrobiological Resources
           Rice Genome Research Program
           2-1-2 Kannondai, Tsukuba
           Ibaraki,
           Japan 305
           Tel: 0298-38-7441
           Fax: 0298-38-7468
           Email: tsasaki@abr.affrc.go.jp
PROJECT   " RGP.
E10189_62.

FEATURES
Source    1..334
           /organism="Oryza sativa"
           /strain="Nipponbare"
           /db_xref="taxon:4530"
           /clone="E10189"
           /clone_1lb="Rice panicle at ripening stage"
           /dev_stage="ripening stage"
           /note="Organ: panicle; Rice cDNA from panicle at ripening
           stage"

BASE COUNT      83 a      76 c      97 g      78 t

ORIGIN
Query Match      11.0%; Score 216.8; DB 108; Length 334;
Best Local Similarity 78.3%; Pred. No. 4.5e-42;
Matches 260; Conservative 0; Mismatches 72; Indels 0; Gaps 0;

QY 1535 gcaaggtgaatcggtgcagaagaagtgtctattaccggtgcgcaacaagtatga 1594
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB   3 GAAAGCGAGATCGGTGAAAGAGCGCTTGGCGCATTCACCGCAGCACCTTCACGA 62

QY 1595 agaacatctaacgtagtactgttagaaacgggaattaaagtacgactagtgtg 1654
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB   63 AGAGCTGTCAACTTCAGAACTGCTGGAACCGGTATCAAGTTATCGACCTGATGTG 122

QY 1655 tccggttcggaaggggttaagtgttatttcggtgtgcggtgtgcggttaaaacct 1714
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB   123 TCCGTTCCGTTAAGGCGGTTAAAGTGTGCTGTGCGTGGTGGGTGTAAGTAAACCGT 182

QY 1715 caatatgtgaattaatcgcgttaacatcgcaattgaagcaactcaggttactctgtcttgc 1774
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB   183 AAACATGATGAGCTCATTCATTCATCGATCGCATCGACATCCGGTTACTCTGTGTTGTC 242

QY 1775 gggggtaagtgagcgtaagcggtgaaggttaacgactctatcatctgtgaagaagactctaa 1834
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB   243 GGGCGTACGTAGTACTCTCGTGAAGGTAAACGACTTTCACGAGAAATGACCGACTCCAA 302

QY 1835 cgtatagataaagtgctctctgtttatgttc 1866
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB   303 CATTATCGACAAAGTATCTCCTGTTATGACC 334

```

```

RESULT 3
LOCUS   BE496766      626 bp      mRNA      EST      03-AUG-2000
DEFINITION BE496766 Hc_d11_08F03_SKPL Haemonchus contortus d11 mixed adult from David
            Knox Haemonchus contortus cDNA clone Hc_d11_08F03 5' similar to
            sp1465611ATPB.CABEL PROBABLE ATP SYNTHASE BETA CHAIN,
            MITOCHONDRIAL PRECURSOR, mRNA sequence.
ACCESSION BE496766
VERSION   BE496766.1  GI:9690101
KEYWORDS EST.
SOURCE    Haemonchus contortus.
ORGANISM  Haemonchus contortus.
           Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabdita; Strongylida;
           Trichostrongyloidea; Haemonchidae; Haemonchinae; Haemonchus.
REFERENCE 1 (bases 1 to 626)
AUTHORS   Blaxter,M.L., Parkinson,J., Whilton,C., Daub,J., Guiliano,D., Hall
           N., Quayle,M. and Barrell,B.
TITLE     Edinburgh University/Sanger Centre Nematode EST Project
JOURNAL   Unpublished (2000)
COMMENT   Contact: Blaxter ML
           Institute of Cell, Animal and Population Biology
           University of Edinburgh
           Ashworth Labs, King's Buildings, West Mains Road, Edinburgh, EH9
           3JF, UK.
           Tel: +44 131 650 6760
           Fax: +44 131 670 5450
           Email: mark.blaxter@ed.ac.uk
           The library was prepared by David Knox of the Moredun Institute,
           Edinburgh. Sequencing was performed by Claire Whilton, ICAPB,
           Edinburgh
           PCR Primers
           FORWARD: T3
           BACKWARD: T7PL
           Plate: 08 row: F column: 03
           Seq primer: SKPL
           High quality sequence stop: 547.
           Location/Qualifiers
           1..626
           /organism="Haemonchus contortus"
           /strain="Moredun"
           /db_xref="taxon:6289"
           /clone="Hc_d11_08F03"
           /clone_1lb="Haemonchus contortus d11 mixed adult from
           David Knox"
           /sex="Mixed"
           /dev_stage="Adult, day 11"
           /note="Vector: Lambda Zap II; Site_1: EcorI; Site_2: XhoI;
           Constructed by David Knox, Moredun Institute, Edinburgh.
           Primary library titre was 4x10^6 pfu."

BASE COUNT      157 a      149 c      179 g      141 t

ORIGIN
Query Match      10.2%; Score 200.8; DB 136; Length 626;
Best Local Similarity 61.4%; Pred. No. 3.6e-38;
Matches 350; Conservative 0; Mismatches 202; Indels 18; Gaps 1;

QY 1351 caggtttgacttgaagttcaacaacaatagtggtgtgtgtgttcgttcgtatcgaa 1410
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB   38 CACGCTTATCTCGAATCTCTCTCAGCATTCAGGACAAATGTGGTCCGTTGCAATGCTA 97

QY 1411 tgggatactcgtatgattaaacgcggtttaagcgtaacaatacgsataaccaattt 1470
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB   98 TGGACGGAACTGAGGGACTTGTTCGGGTCAGAAAGTTGTTGACACGTGATGCAATCA 157

QY 1471 ctgttcaagtggaacggaacaatctgttcgtatcatgaagttatgttggtgaaccaatcg 1530
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB   158 AGATTCCAGTAGGCGCAGAGACTCTCGGCGCAATATGATGATGATCGGCGCAACCAATCG 217

QY 1531 atgaagcaagtgtaaatcgtgtgcagaagaagtgtctattcaaccgtgcgacccaatt 1590
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB   218 ACGAGCGAGGTCCTATCCCGCCAGCACCGCGCTCTATTTCACCTGAGGACCAAGAT 277

```


[illegible]

RESULT	4
LOCUS	A1975327
DEFINITION	A1975327 606 bp mRNA EST 27-AUG-1999
ACCESSION	EST269921 Schistosoma mansoni female, Phil Love/Steve Merrick
VERSION	Schistosoma mansoni cDNA clone SMR46 5' end, mRNA sequence.
KEYWORDS	A1975327.1 GI:5786495
SOURCE	EST.
ORGANISM	Schistosoma mansoni.

REFERENCES

AUTHORS
Merrick, J. M., Osman, A., Loverde, P. T., Chandra, I., Glodek, A., Fraser,
C. M. and Lee, N. H.

JOURNAL Unpublished (1998)
COMMENT Contact: Norman H. Lee

The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel.: 301 838 3529
Fax: 301 838 0208
Email: nhleeth@igf.org
Seq primer: M13 Reverse.

FEATURES	Location/Qualifiers
source	1. .606

```

/organism="Schistosoma mansoni"
/db_xref="taxon:6183"
/clone="SMFAR46"
/clone_id="Schistosoma mansoni female, Phil Laverde/Jo
Merrick"
/sex="female"
/note="Vector: pluescript SK; Site_1: EcoRI; Site_2: XhoI
; directionally cloned cDNAs"
164 a 121 c 159 g 162 t

```

Query Match	10.1%	Score 199;	DB 104,	Length 606;
Best Local Similarity	60.4%;	Pred. No. 9.7e-38;		
Matches 356;	Conservative 0;	Mismatches 215;	Indels 18;	Gaps 1,

OY 1354 gttttagctactgaagttccaacaacattagtgtgtgtatgctgtatcgcaatbg 1413
||| ||||| | ||||| | ||||| ||| ||| ||| |||
Db 18 GTCCTTATACCTGAGATCGCACACAATCTTGCGAGAACACTGTGGCACCATTGCCTATGG 77

Qy 1414 gatcatctgatgattaaacgcggttaagcgttaacaataacgaataacccaattctg 1473
 - - - - -
 Db 78 ACCGTACCGAAGGTCTAGTTCGGCGGCACGACAGTGTGTTGATACCTGGCGGTCCATTTCGA 137

QY 1474 ttccagctgggaacgaacatttggctgcatactgaacgtatttgggtgaaccaatcatg 153
||||| - - - - - ||||| - - - - - ||||| - - - - - ||||| - - - - - |||||
Db 138 TTCTCTGTTGGTCCCTGGAACCCCTTGCCCGTTCATGAACTGTATCGAGAACCAATCGATG 197

OY 1534 agcaagtgaaatcgtgtcgaagaagaattgctctactcacgttgcgccaccaagttag 159
||| ||| | | | | | | | | |
Db 198 AGAAGAGGCCCATCAAAACGAAAATGATGTCTGGGTATTACCAAGAGTGTCCAGACTTCA 257

OY 1594 aagacacatctaacytactgaaccttttagaacggyaaattaagtatacgacttagtt 1655
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Dd 258 TTGAAATGAGTACTCGACCAGAATAATTTCAGACAGGCATCAAGGTGTGAAGCCTGTAG 317

DY 1654 gtccggttcgcgaagggcgtaagaatgattatcttgctgctggtcgcgttaaaccg 1713
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 318 ctcttattatgcgaaggcgaataaatcgcttttttcgctgctgcctggcgtttgccaaactcg 377

DY 1714 tcaatgatgagtaattacatccgtacaatcgaattgagaccacagttactctgttc 1773
| | | | | | | | | | | | | | | | | | | | | |
Db 378 TACTTATTATGAGACCTAATAATAATTGCCAGACCTACAGGAGTTACTGTGGTTG 437

QY 1774 cgggggfaggtlgagcgcgcfgaaggtfaagacattcatgatgaagactcta 1833

Db 438 CTGCTGTTGGTGAACGTAACTCGGAAGGCAATATTATATACAGAAATGAATACAACTG 497

Dy 1834 acgatttaga-----faangtgcctcgtttatgtccaatgaacg 1875
||| |||
Dd 498 GTGTGATTGAACCTTAAGGAACAATTCATAAGTGTCTCGTGATCAAGGGCAAAGAATFG 557

QY 1876 agccaccaggtaacggtttacgtgtgcaattacagcgttaactatggc 1924
+ ||||| ||| ||||| || |||| || |||| || |||| || ||||
Db 558 AACCAACAGGAGCACGCTCGTGTGGCTTAACGGGCTTAACAATTC 606

RESULT	5
AV404766	
AV104766	
600 bp	
mpm	
pcr	
06-Feb-2000	

LOCUS	CHROM	START	END	STRAND	FEATURES
AV404766	1	1039	1040	+	UTR
AV404766	1	1040	1041	+	UTR
AV404766	1	1041	1042	+	UTR
AV404766	1	1042	1043	+	UTR
AV404766	1	1043	1044	+	UTR
AV404766	1	1044	1045	+	UTR
AV404766	1	1045	1046	+	UTR
AV404766	1	1046	1047	+	UTR
AV404766	1	1047	1048	+	UTR
AV404766	1	1048	1049	+	UTR
AV404766	1	1049	1050	+	UTR
AV404766	1	1050	1051	+	UTR
AV404766	1	1051	1052	+	UTR
AV404766	1	1052	1053	+	UTR
AV404766	1	1053	1054	+	UTR
AV404766	1	1054	1055	+	UTR
AV404766	1	1055	1056	+	UTR
AV404766	1	1056	1057	+	UTR
AV404766	1	1057	1058	+	UTR
AV404766	1	1058	1059	+	UTR
AV404766	1	1059	1060	+	UTR
AV404766	1	1060	1061	+	UTR
AV404766	1	1061	1062	+	UTR
AV404766	1	1062	1063	+	UTR
AV404766	1	1063	1064	+	UTR
AV404766	1	1064	1065	+	UTR
AV404766	1	1065	1066	+	UTR
AV404766	1	1066	1067	+	UTR
AV404766	1	1067	1068	+	UTR
AV404766	1	1068	1069	+	UTR
AV404766	1	1069	1070	+	UTR
AV404766	1	1070	1071	+	UTR
AV404766	1	1071	1072	+	UTR
AV404766	1	1072	1073	+	UTR
AV404766	1	1073	1074	+	UTR
AV404766	1	1074	1075	+	UTR
AV404766	1	1075	1076	+	UTR
AV404766	1	1076	1077	+	UTR
AV404766	1	1077	1078	+	UTR
AV404766	1	1078	1079	+	UTR
AV404766	1	1079	1080	+	UTR
AV404766	1	1080	1081	+	UTR
AV404766	1	1081	1082	+	UTR
AV404766	1	1082	1083	+	UTR
AV404766	1	1083	1084	+	UTR
AV404766	1	1084	1085	+	UTR
AV404766	1	1085	1086	+	UTR
AV404766	1	1086	1087	+	UTR
AV404766	1	1087	1088	+	UTR
AV404766	1	1088	1089	+	UTR
AV404766	1	1089	1090	+	UTR
AV404766	1	1090	1091	+	UTR
AV404766	1	1091	1092	+	UTR
AV404766	1	1092	1093	+	UTR
AV404766	1	1093	1094	+	UTR
AV404766	1	1094	1095	+	UTR
AV404766	1	1095	1096	+	UTR
AV404766	1	1096	1097	+	UTR
AV404766	1	1097	1098	+	UTR
AV404766	1	1098	1099	+	UTR
AV404766	1	1099	1100	+	UTR
AV404766	1	1100	1101	+	UTR
AV404766	1	1101	1102	+	UTR
AV404766	1	1102	1103	+	UTR
AV404766	1	1103	1104	+	UTR
AV404766	1	1104	1105	+	UTR

VERSION	AV404/00.1	GI.0306034
KEYWORDS	EST.	
SOURCE	domestic silkworm.	
ORGANISM	Bombyx mori	
	Phyllosticta; Nectria; Arthrospira; Trachospora; Dothidea; Tenebrio;	

REFERENCE

Mita K., Morimoto M., Shimada T., Okano T. and Maeda S.

AUTHORS: Mita, K., Morimoto, M., Shimada, T., Ohtsu, H., and Maeda, S.
 TITLE: BcdyX mori cDNA
 JOURNAL: Unpublished (2000)
 COMMENT: Contact: Mita K
 Genome Database Group

Genome Research Group
National Institute of Radiological Sciences
Anagawa 4-9-1, Inage, Chiba 263-8555, Japan
Email: kmita@nirs.go.jp
method:uni-directional, sequence direction:sequenced from T3 pro
(5' -> 3')
Project='Silksworm Genome Program in MAF, and Research for the
Future Program in JSPS', see 'SilkBase',
<<http://www.ab.a.u-tokyo.ac.jp/silkbase/>>, for whole ESTdb.

```

FEATURES
  source
    location/Qualifiers
      1. .699
        /organism="Bombyx mori"
        /strain="p50"
        /db_xref="taxon:7091"
        /clone="pry0416"
        /clone_1lb="Bombyx mori prothoracic gland p50 5th-instar
day-4 larva"
        /sex="female/male mixed"

```


SOURCE domestic silkworm.
ORGANISM Bombyx mori
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia; Bombycoidea; Bombycidae; Bombyx.
REFERENCE 1 (bases 1 to 746)
AUTHORS Mita,K., Morimyo,M., Shimada,T., Okano,K. and Maeda,S.
TITLE Establishment of cDNA database of Bombyx mori
JOURNAL Unpublished (1999)
COMMENT Contact: Mita K
Genome Research Group
National Institute of Radiological Sciences
Anagawa 4-9-1, Inage, Chiba 263-8555, Japan
Email: kmits@nirs.go.jp
PROJECT - 'CREST project by JST',
Location/Qualifiers
1. 746
/organism="Bombyx mori"
/strain="p50(Daizo)"
/db_xref="taxon:7091"
/clone="e40836"
/clone_lib="Bombyx mori p50(Daizo)"
BASE COUNT 191 a 159 c 202 g 194 t
ORIGIN

Query Match 9.5%; Score 187.6; DB 106; Length 746;
Best Local Similarity 57.5%; Pred. No. 5.7e-35;
Matches 359; Conservative 0; Mismatches 259; Indels 6; Gaps 1;

OY 1220 cagcgcaatttaacaaatagagatcggttaagcgaactgggtaaaattgtacaatc 1279
DB 109 CAGAGGAGACTTGTGCTCAATTCGACGAAAGCCCAAGTTAGTGTTCCTTAT 168
OY 1280 cggcggttattgacgttgaatcccaaga-----tgcaagcacaagaatata 1333
DB 169 CGGTGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 228
OY 1334 tgccttaattgtgaacaggtttgacttgaattcaacaaattgtgtgtgt 1393
DB 229 GGTGCAAAATCATCTCCCGCTGCTGATGATGATGATGATGATGATGATGAT 288
OY 1394 agtgcgcatcgcgaatggatcgcgtgatgattaaacggttgaagcgaacaa 1453
DB 289 ACTTGGACACTTGGCATGACGATGATGATGATGATGATGATGATGATGATGAT 348
OY 1454 taagaataaccgaatttcgttcagtggaacgaacaaattgtgtgtgtatgaacgt 1513
DB 349 CTCTGGCTACCAATTCGATCCGCTGAGGACTGAAACCTCGGTGCATCATCAT 408
OY 1514 attggtgaacaaatcgatgagcaagtgaatcggttcagagaagaattgtctattca 1573
DB 409 AATCGGCAACCGATTGACGAGCGGTGCCATGCCACGCAAGACGTGCTATTGA 468
OY 1574 ccgttcgcacccaagttaagaacaatctaacagttactgaactttgaagaggaat 1633
DB 469 TGCTAAGCTCCAGAGTTTGTGACATGCTGTGACAGCAGGAGATTCTGTAACT 528
OY 1634 taaagttaacgaactagtttcgttcgttcgagaaggggttaagaattgttcgttgg 1693
DB 529 AAAATGCTGCTGCTGCTGCTGCTTATGCCAAAGAGGAAGATTGGGTGTCGG 588
OY 1694 tggcggttcggttaaacgcgtcaataatgtaataatcgttaacatcgcaattgagga 1753
DB 589 AGCTGCTGAGGCAAACTGATTTGATTAATGAACTGAACTGAACTGTTGCAAG 648
OY 1754 ctcaagttactctgttcgttcggttcggttcggttcggttcggttcggttcggttc 1813
DB 649 TGCTGCTACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 708
OY 1814 tcatgatgaagaactcaact 1837
DB 709 CCACGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 732

RESULT 8
AL533571 982 bp mRNA EST 13-FEB-2001
LOCUS AL533571 LTI_FL013_Fbrn1 Homo sapiens cDNA clone CS0DF002Y109 5
DEFINITION prime, mRNA sequence.
ACCESSION AL533571
VERSION AL533571.1 GI:12797064
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 982)
AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
Location/Qualifiers
1. 982
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DF002Y109"
/clone_lib="LTI_FL013_Fbrn1"
/dev_stage="pooled tissue from post conception fetuses (20 week, 24 week and 26 week)"
/lab_host="DH10B"
/note="organ: Fetal brain; Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was constructed by life technologies. Contact : Feng Liang Life Technologies, a division of invitrogen 9800 Medical Center Drive Rockville , Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifestech.com URL : http://fulllength.invitrogen.com"

BASE COUNT 224 a 223 c 281 g 249 t 5 others
ORIGIN

Query Match 9.4%; Score 185.8; DB 106; Length 982;
Best Local Similarity 55.0%; Pred. No. 1.6e-34;
Matches 389; Conservative 4; Mismatches 296; Indels 18; Gaps 1;

OY 1284 gcggttattgacgttgaattcccaagaatgcagtaaccgaagatatatgaccttaatt 1343
DB 207 GTGGACGTCCACTTTGATGAGGAGGACTACCAATTCCTAAATGCCCTGGAAGTC 266
OY 1344 gtgaacaggttgaattgaattcaacaaattgtgtgtgtgtgtgtgtgtgtgtgt 1403
DB 267 AGGAGACCAAGCTGTTTGTGAGGTGGCCAGCATTTGGGTGAGAGCACAATAAGAT 326
OY 1404 atcgcaatggatcactcgtatgattaaacgcggtttaaagcgttaacaaatgaataac 1463
DB 327 AATGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 386
OY 1464 ccaatttcgttcagtggaacgaacaaattgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1523
DB 387 CCAATCAAAATTCCTGTTGCTCTGATTTGGGCGAGAAATGATGATGATGATGAT 446
OY 1524 ccaatcgatgagcaaggtgaatcggttcgagaagaattgtgtgtgtgtgtgtgtgtgt 1583
DB 447 CCTATTGATGAAGAGRGTCCTCAATCAAAACCAATTTGCTCCATTTATGCTGAGGT 506
OY 1584 ccaagttatgaagaacaaatcaacagttactgaactttgaagaacgggaattatc 1643
DB 507 CCAGAGTTCATGAAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 566

RESULT	9
AL514744	
LOCUS	
DEFINITION	AL514744 918 bp mRNA EST 13-FEB-2001
ACCESSION	AL514744 LTI.NF0006.Pl2 Homo sapiens cDNA clone CLOBB013ZH09 5
VERSION	prime. mRNA sequence.
KEYWORDS	AL514744
SOURCE	AL514744.1 GI:12778237
ORGANISM	EST.
	human.
	Homo sapiens
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE	1 (bases 1 to 918)
AUTHORS	Li, W.B., Gruber, C., Jesssee, J. and Polayes, D.
TITLE	Full-length cDNA libraries and normalization
JOURNAL	Unpublished (2001)
COMMENT	Contact: Genoscope
	Genoscope - Centre National de Sequencage
	BP 191 91006 EVRY cedex - France
	Email: seqref@genoscope.cns.fr, web : www.genoscope.cns.fr.

Query Match	9.48;	Score 184.6;	DB 105;	Length 918;
Best Local Similarity	55.08;	Pred. No. 3.1e-34;		
Matches 389;	Conservative	3;	Mismatches 297;	Indels 18;
				Gaps 1;

Oy 1284 gcggttattgacgtttgaattccccaagaatgcagtlaccaaaagtatatgatgctctaata 1343
| | | | |
Db 206 GTGGACGCTCCAGTTTGATGAGGAGACTCMACCAAATTCAATAATGCCCTGGAACTGCAGAAGC 265

RESULT	10
AL518359	
LOCUS	978 bp mRNA EST 13-FEB-2001
DEFINITION	AL518359 LTL.NFL011.NBC1 Homo sapiens CDNA clone CSDDA009YB03 5 prime, mRNA sequence.
ACCESSION	AL518359
VERSION	AL518359.1 GI:12781852
KEYWORDS	EST.
SOURCE	human.
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
TITLE	1 (bases 1 to 978)
JOURNAL	Li,W.B., Gruber,C., Jesse,J. and Polayes,D.
COMMENT	Full-length cDNA libraries and normalization Unpublished (2001) Contact: Genoscope National de Semeurage

CONTACT: genoscope
Genoscope - Centre National de Séquençage
BP 191 91006 EVRY cedex - France
Email: secretairegenoscope.cns.fr, Web : www.genoscope.cns.fr
Location/Qualifiers
1
c170

```
FEATURES
source
1..978
location/Qualifiers
/organism="Homo Sapiens"
/db_xref="taxon:9606"
/clone="CS0DA0091B03"
/clone_id="LTI_NFL011_NB01"
/sex="male"
```


QY	1728	ttatctcgttaacctcgtcaattgagacaccagttacctctgtcttcgtcggggtagtgag	1787
Db	650	TTTATTCACACATNTCCGCAAAAGCCCATGGTGGTTACTGTGCTTGCTGGGTGGTGAG	709
QY	1788	cgtacgcgttgaagttacgactcttatacatalgatalgaagaacctaacgtataga----	1843
Db	710	AGGACCCGGTAAAGGCATATTTATTCATCTCAATGAATGATGATCTGGTATTAACCTTA	769
QY	1844	-----taagtgtctctgttctatgtgtcaatgaacgaacccaacagttaac	1889
Db	770	AAGAGTCCACCTCTTAAGTAGTAGCGCTGGTATATGTGTCAAATATAAAGAACACACTGGTGCT	829
QY	1890	cgattacgttggcgcattaaacagccttaactatgtgcggaaaaattccgtgatgaagtgct	1949
Db	830	CGTCCCGGGTACTCTGTACTGGGCTGACTGTGGCTGATTAACCTTACAGACCAAGAAAGT	889
QY	1950	gagtcctatcttcgttgtaaa	1972
Db	890	CAAGATGTAAGTATTTATTTGA	912

RESULT	12
AL536748	
LOCUS	931 bp mRNA EST 13-FEB-2001
DEFINITION	AL536748 L1L_FL013_Fbrn1 Homo sapiens cDNA clone CSDDFC038YK02 5 prime, mRNA sequence.
ACCESSION	AL536748
VERSION	AL536748.1 GI:12800241
KEYWORDS	EST.
SOURCE	human.
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
TITLE	1 (Bases 1 to 931)
JOURNAL	I.I.W.B., Gruber,C., Jesse,J. and Polayes,D.
COMMENT	Full-length cDNA libraries and normalization unpublished (2001)
Contact:	Gemoscope "attf@scg.univ-paris1.fr"

Genoscope - Centre National de Séquençage
BP 191 91006 EVRY cedex - France
Email: segrete@genoscope.cns.fr, Web : www.genoscope.cns.fr
location/Qualifiers
1. .931

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CSDF038YK02"
/clone_id="Jrl_FL013_FBrn1"
/dev_stage="pooled tissue from post conception fetuses (20
week, 24 week and 26 week)"
/lab_host="DH108"
/note="Organ: Fetal brain; Vector: PCWSPORT 6; 1st strand
cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-stranded cDNA was digested with Not I
and cloned into the Not I and Eco RV sites of the
PCWSPORT 6 vector. Library was constructed by Life
Technologies. Contact : Feng Liang Life Technologies, a
division of Invitrogen 9800 Medical Center Drive Rockville
, Maryland 20850, USA Fax : (1) 301 610 8371 Email :
fliang@lifestech.com URL :
http://fulllength.invitrogen.com"

```

Query Match	9.3%	Score 18.8	DB 106	Length 931
Best Local Similarity	5.2%	Pred. No. 4.9e-34		
Matches 390	Conservative 1	Mismatches 299	Indels 18	Gaps 1
Oy	1284	gcgcgttatcgcgttggaattcccaagaagtcagcagccaaagtatatagatcccttaaat	1343	
db	207	GTGGAGCGTCCATCTTGGATGATGGAGCTACCAACCAATTTCTAAATGCCCTGGAATGCCAAGCC	266	

OY	1344	gttgaacaggttttagtacttgatgaagttcaacacacattgagtggtgtgtatgcgcgt	1403
Db	267	AGGGAGACCCAGCACTGGTTTGGAGGTGGCCACGACATTTGGTGAGACACAGTAAGGACT	326
OY	1404	atcgcaatgagatcatctgtatgattaaacgcggttlaagcgtlaacaatacgaataac	1463
Db	327	ATTGCTATGAGATGGTACAGAGAGCTTGTTAGAGGCCAGAAAGTACTGATTCGTGTGCA	386
OY	1464	ccaatttcgtccagtggtgggaacgaaacattgggtcgtatcatatgaacgtatgggttaa	1523
Db	387	CCAATCAAAATTCCTGTTGGTCTCTGAGACTTTGGGAGATCATGATCATTTGGAGAA	446
OY	1524	ccaatcgatagaagaagttgaacgtgtcgtagcagaagaatttgcataatccgttcgcga	1583
Db	447	CCATTATATGAAGAAGGTCCCATCAACCAACAAACATTTTGCTCCATTCATGCTGAGGCT	506
OY	1584	ccaagttatgaagacacatctaacagbaacttgaacctlttgaacagggaaattaaagttac	1643
Db	507	CCAGAGTTCCATGAGAAATGAGTGTGAGACAGAAATTCGTGTGACTGGTATCAAGGTTGTC	566
OY	1644	gaacttagttgtccgttcttggaagaagggtgaagaagtgtatctcgtagtgcggtgc	1703
Db	567	GATGTGCTAGCTCCCTATGCCCAGAGGGGCGCAAAATTGKCTTTTGTGTGCTGGAGTT	626
OY	1704	ggtaaacgcgtcaatgatgatgaatlaatccgtlaacatgcgaattgaacctgaagttac	1763
Db	627	GGCAAGACTGTACGATGATCATGATGAGATTATCAACAATGTCGCCAAAGCCCATGTGCTTAC	686
OY	1764	tctgtccttgcgggggtaggttgagcgtgaacgctggaagtgtaagcaattctatctatgaag	1823
Db	687	TCTGTGTTTGGTGGTGTGTGTGAGAGAGACCCGTGAAGGCATATGTTTATACATGAAATG	746
OY	1824	aaagacctaacgctattaga-----taagtgctccttgittatggt	1865
Db	747	ATTGATCTGGTGTATACACTTAAAGATGCGCACCTTAAGGATGCGGTGATATATGT	806
OY	1866	caaatgaacagagaccacaggtaacgcgttactgctgtgcgcatataacaggtctaactatggcg	1925
Db	807	CAAAATGATGAACACCACTGCTGTCCTGTCGCCCGGGGTACTCTGATGGGCTAGCTGTGGCT	866
OY	1926	gaaataactcgtatgaaggtcgtgatgtcttactcttcgttgataa	1972
Db	867	GAAATCTTCAGAGACCAAGAAGCTCAACATGTACTGCTATTATTATGA	913

RESULT	13
LOCUS	AL516904
DEFINITION	AL516904 985 bp mRNA EST 13-FEB-2001
ACCESSION	AL516904
VERSION	AL516904
KEYWORDS	prime, mRNA sequence.
SOURCE	AL516904.1 GI:12780397
ORGANISM	EST. human.
REFERENCE	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
AUTHORS	1 (bases 1 to 985)
TITLE	I.L.W.B., Gruber,C., Jesse,J. and Polayes,D.
JOURNAL	Full-length cDNA libraries and normalization
COMMENT	Unpublished (2001) Contact: Genoscope

```

FEATURES
source
location/Qualifiers
1..985
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0D4N07YG13"
/clone_id="LTI_NFL011.NBC1"
/sex="male"
/tissue_type="neuroblastoma cells"

```



```

Db 633 GGCAGAGCTGCTACTGATCATGAGATGATCAACATGTCCGCCAAGCCCATGTGTATAC 692
QY 1764 tctgtcttgcgggggtaagtgagtcgagcgtggaagtgtaacgactctctatgaatg 1823
Db 693 TCTGTGTTGCTGCTGCTGTGGTGAGAGAGACCCGTGAAGGCAATGTTATACCATGAATG 752
QY 1824 aaagactctaacgtattaga-----taagtgtctctgtttatcgt 1865
Db 753 ATTGAATCTGCTGTATCAACTTAAAGATGCCACCTTAAGTAGCGGTATATGCT 812
QY 1866 caaatgaacgagccacccaggttaacggttaccgtgtgtgcatgaacaggcttaactatg 1925
Db 813 CAATGATGATGACACCTGCGTCTGCGCCGGGTAGCTGACTGAGCTGAGCTGAGTGGCT 872
QY 1926 gaaaaattccgtgatgaagtcgagtcgtatcttctcgttgaata 1972
Db 873 GAATACTTCAGAGACCAAGATGCAAGATGCTACTGCTATTATTGA 919

```

```

RESULT 15
AL514536 1047 bp mRNA EST 13-FEB-2001
LOCUS AL514536 LRI_NFL006_PL2 Homo sapiens cDNA clone CL08B003ZC04 5
DEFINITION prime, mRNA sequence.
ACCESSION AL514536
VERSION AL514536.1 GI:12778030
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1047)
AUTHORS I., W. B., Gruber, C., Jessup, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES
Source
1.1047
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CL08B003ZC04"
/clone.lib="LRI_NFL006_PL2"
/tissue_type="placenta"
/note="Vector: pCMVSPORT 6, Site_1: NotI, 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library was normalized. Library was constructed by
Life Technologies. Contact : Feng Liang Life Technologies,
a division of Invitrogen 9800 Medical Center Drive
Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com"
BASE COUNT 239 a 237 c 302 g 263 t 6 others
ORIGIN

```

```

Query Match 9.3%; Score 183.4; DB 105; Length 1047;
Best Local Similarity 55.0%; Pred. No. 6.2e-34;
Matches 389; Conservative 2; Mismatches 298; Indels 18; Gaps 1;

```

```

QY 1284 gcggtatgacgttgatccacagatgcagtaacaaagtatatgaccttaaat 1343
Db 207 GTGGAGCTCCAGTTGATGAGGAGGACTACCAATCTTAATGCCCTGGAAGTGCAAGGC 266
QY 1344 gttgaacaggttgatgacttaagttcaacaaacaaatgagtggtgtgtgtagtcgtg 1403
Db 267 AGGAGACCAAGCTGTTGGAGGTSGCCAGCATTTGGTGAGAGACACAGTAAGGACT 326

```

```

QY 1404 atcgcaatggatcatctgtgtgattaaacgcggtttaagcgttaacaaatacgaataac 1463
Db 327 ATTGCTATGATGGATGATACAGAGCTGCTTAGAGGCCAGAAAGTACTGGAATTTCTGGTGA 386
QY 1464 ccaattctgttccagtgagggaacgaaacattggtcgtatcactgaacgatttggatgaa 1523
Db 387 CCATTCAAATTCCTGTTGGTCCCGACACTTGGGCGAATTCATGAAATGATCATTTGAGAA 446
QY 1524 ccaatcgtatgacgaagtgaaatcgtgtgcagaagagaattggtctatcaccgtgcga 1583
Db 447 CCTATGATGAAGAGAGTCCCATCAAAACCAAAATTCCTCCCATTCATGCTGAGGCT 506
QY 1584 ccaagttatgaagacacatctaaacagtaacttttagaagaagggaattaaagtatc 1643
Db 507 CCAGAGTTCATGTAAGTAAGAGTGTGACGACGAAATTCGTGACTGCTATCAAGCTTGTGTC 566
QY 1644 gacttagttgtccgttgcgaaagggtaaagttagttatcgttgggtgcggtgtc 1703
Db 567 GATCTGCTAGCTCCCTATGACCAAGGTGCAAAATTTGGGCTTTTGGTGGTGGTGAAGTT 626
QY 1704 ggtaaacccgtcaatataatgaaatlaattccgttaacatcgcaattgagcaactcaat 1763
Db 627 GGCAGAGCTGTTCTGATCATGAGTATCAATGATGCGCAAAAGCCCATGCTGTGTATAC 686
QY 1764 tctgtcttgcgggggtaagtgagcgtgagcgtgaaaggttaacgactctatcatgagatg 1823
Db 687 TCTGTGTTGCTGCTGTTGGTGAGAGAGACCCGTGAAGGCAATGATTTATACATGAATG 746
QY 1824 aaagactctaacgtattaga-----taagtgtctctgtttatcgt 1865
Db 747 ATTGAATCTGCTGTATCAACTTAAAGATGCCACCTTAAGTAGCGCTGATATGCT 806
QY 1866 caaatgaacgagccacccaggttaacggttaccgtgtgtgcatgaacaggcttaactatg 1925
Db 807 CAATGATGATGACACCTGCGTCTGCGCCGGGTAGCTGACTGAGCTGAGCTGAGTGGCT 866
QY 1926 gaaaaattccgtgatgaagtcgagtcgtatcttctcgttgaata 1972
Db 867 GAATACTTCAGAGACCAAGATGCAAGATGCTACTGCTATTATTGA 913

```

Search completed: October 28, 2001, 16:44:14
Job time: 6333 sec

10/31/2001 13:40:41

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 28, 2001, 16:35:36 ; Search time 74.11 seconds
(without alignments)
5037.398 Million cell updates/sec

Title: US-09-545-199C-3
Perfect score: 1972
Sequence: 1 agcggccattgctcagc.....gtctattctctgtgataa 1972

Scoring table:
OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 324599 seqs, 9465562 residues

Word size : 25

Total number of hits satisfying chosen parameters: 1

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 1000 summaries

Database : Issued Patents_NA:*

1: /cgn2_6/ptodata/1/ina/5A.COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B.COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A.COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B.COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCITUS.COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	44	2.2	1374	2	US-08-743-637B-187 Sequence 187, App

ALIGNMENTS

RESULT 1
US-08-743-637B-187
Sequence 187, Application US/08743637B
Patent No. 5694066

GENERAL INFORMATION:

APPLICANT: BERGERON, Michel G.
APPLICANT: PICARD, Francois J.
APPLICANT: OUELLETTE, Marc
APPLICANT: ROY, Paul H.
TITLE OF INVENTION: SPECIES-SPECIFIC AND UNIVERSAL DNA
TITLE OF INVENTION: PROBES AND AMPLIFICATION PRIMERS TO RAPIDLY DETECT AND
IDENTIFY COMMON BACTERIAL PATHOGENS AND ASSOCIATED
TITLE OF INVENTION: ANTIBIOTIC RESISTANCE GENES FROM CLINICAL SPECIMENS ...
NUMBER OF SEQUENCES: 273
CORRESPONDENCE ADDRESS:
ADDRESSEE: QUARLES & BRADY
STREET: 411 EAST WISCONSIN AVENUE
CITY: MILWAUKEE
STATE: WISCONSIN
COUNTRY: USA
ZIP: 53202-4497

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/743,637B
FILING DATE: 04-NOV-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/526,840
FILING DATE: 11-SEP-1995
ATTORNEY/AGENT INFORMATION:
NAME: BAKER, Jean C.
REGISTRATION NUMBER: 35,433
REFERENCE/DOCKET NUMBER: 850586.90012
TELECOMMUNICATION INFORMATION:
TELEPHONE: (414) 277-5000
TELEFAX: (414) 277-5591
INFORMATION FOR SEQ ID NO: 187:
SEQUENCE CHARACTERISTICS:
LENGTH: 1374 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: Haemophilus influenzae
US-08-743-637B-187

Query Match 2.2%, Score 44, DB 2, Length 1374;
Best Local Similarity 100.0%, Pred. No. 2.1e-12;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 607 GTTATGTCGAATGACGACGACGAGTACCGTTTACGCT 650
OY 1857 gttaatgcaatgaacgagccaccaggttaaccgttacgtgt 1900
|||||

Search completed: October 28, 2001, 17:56:16
Job time: 4840 sec

